INTERMEDIATE PHYSICS

INTERMEDIATE PHYSICS

WILEY EASTERN PRIVATE LIMITED
NEW DELHI

First Printed, 1932 ; Reprinted 1933, 1934
Second Edition, 1935
Reprinted, 1936, 1937, 1938, 1940, 1941, 1942, 1943 (twice), 1944, 1945
Third Edition, 1947

PREFACE

In this book an endeavour has been made to cover the syllabuses required in physics for the Intermediate, the Higher School Certificate, and Scholarship Examinations of the various universities. Although the selection of the material which is to appear is in some measure a matter of personal taste, and other teachers may assess the various parts of the subject differently, it is hoped that a study of the following pages will furnish the student with a comprehensive knowledge of the essential principles of elementary physics, and provide him with a useful tool for further work in this and other subjects. It is this latter aspect which accounts for the somewhat numerous references to the applied sciences. In order thus to equip the student and complete the argument as far as space and the mathematical attainments of the student would permit, the author has not hesitated to use the calculus notation and, in one or two instances, the powerful and beautiful methods of the calculus itself. In this way it is hoped that students will acquire, in the earlier stages of their careers, knowledge which is essential if they are properly to appreciate the aims of physics, and moreover, knowledge which must be possessed before work for a degree in physics is attempted. The author firmly believes that such knowledge must be attained at an early stage if the task of the student in mastering the more advanced parts of his subject is not to be too arduous.

In presenting this third edition to his readers, the author has kept in view three chief aims, viz. (i) to explain in greater detail the more elementary parts of the subject, as well as those parts which usually appear difficult on a first acquaintance with them; (ii) to add an account of that portion of physics essential to scholarship candidates and to those who desire to obtain more than a superficial knowledge of the subject; (iii) to endeavour to give definitions and to use equations which are correct dimensionally. Usually, in an elementary exposition of physics, the dimensions of a physical quantity are not considered—a course leading to much trouble in later years. In order to indicate those parts of the book which are generally considered to be rather above Intermediate standard, they have been printed in smaller

type: such portions should certainly be omitted on a first reading.

In Part I there is a general account of the properties of matter where the subject of surface tension has been treated on the basis of the idea of surface energy, i.e. molecular happenings in the liquid itself. The subjects of diffusion, osmosis, and elasticity, have been treated in a somewhat detailed manner. A brief account of the theory of dimensions and examples of its use have been added.

In Part II an elementary exposition of the subject of heat is presented. Here the author has endeavoured to give brief accounts of some of the more modern and accurate methods of obtaining thermodynamical data. In particular, the fundamental principles of continuous-flow calorimetry have been developed, and the method then applied to the determination of latent heats of vaporization and thermal conductivities. In order to maintain uniformity with the rest of the book, the specific heat of a substance has been defined in such a way that its dimensions are, in one system of units, cal. gm.$^{-1}$ deg.$^{-1}$ C. This seems desirable, since the dimensions of all equations appearing in the subject of heat are then correct. The chapter on thermal conductivity remains greatly extended; attention is there directed to the distribution of temperature in bars along which heat flows under different conditions; a brief account of modern guard-ring methods is given as well as an application of this method to liquids. In the chapter on the first law of thermodynamics the historical development is emphasized. Some parts of the chapter on radiation have been rewritten, the development now being more logical: an effort is also made to draw a clear distinction between processes depending only on the emission or absorption of radiant energy, and those in which the processes of radiation, conduction and convection are simultaneously involved. Moreover, the determination of specific heats by the method of cooling is described in the chapter on calorimetry—not as usual, following an account of Newton's law of cooling, for the method is independent of the validity of this law.

Optics forms the subject of Part III, and here an effort has been made to expound the principles of tracing rays through an optical system; it is only by the actual carrying out of such tracings that a thorough acquaintance with the elementary principles of optical instruments may be obtained. In dealing with the subject of magnification, this has been regarded as a numerical quantity, so that any formulæ for magnification only contain positive entities. These are denoted in the usual manner by $|x|$, etc. Students seem to find this method the least difficult of all. The chapter on optical instruments has been rewritten and considerably extended: the

treatment is now up to date. The subjects of interference, diffraction, and polarization are treated in still greater detail. The theoretical part has been made to depend on ideas involving the time of transit between two points rather than on the number of waves.

In Part IV there follows a brief survey of acoustics, where a short account of the modern methods of sound-ranging on land and sea has been given. Here there appears a comprehensive account of methods for determining the velocity of sound in air and in sea-water: a short section on supersonics is included. The treatment of Lissajou's figures is now more complete.

Part V, that section of the book dealing with electricity and magnetism, has been considerably extended. This section begins with an account of electrostatics, in which there is included a chapter on the theory of isotropic dielectrics. Here the idea of 'electric displacement' is developed and a brief account of Debye's work on the dielectric constants of gases follows. Gauss's theorem and its applications are then discussed. Electrostatic instruments are treated quite fully and in an up-to-date manner. A section on magnetism follows: here, as in electrostatics, the term 'strength of field' is generally used in preference to 'intensity of field'. The symbol H now denotes the strength of a magnetic field, so that another symbol, e.g. H_0, must be selected to denote the horizontal component of the earth's magnetic field. In this section on magnetism there is given a short account of an elementary form of the Schuster magnetometer and of instruments used for recording continuously variations in the magnetic elements. In the opening remarks of the first chapter on current electricity, the connexion between electricity produced by friction and voltaic electricity is discussed. Many changes appear in the succeeding chapters where a fairly full account of accurate methods of measuring a current and a resistance has been given. The underlying ideas have then been applied to the determination of small resistances and of small potential differences. More attention has been given to the design and principles of construction of electrical measuring instruments. The chapter on the magnetic properties of iron and steel has again been enlarged: it includes a brief discussion of paramagnetic and diamagnetic substances. The chapter on electromagnetic induction has been thoroughly revised. Here, as in other parts of the book, greater stress has been laid on historical facts, the pioneer work of Faraday being followed step by step. Many new diagrams showing the lines of magnetic induction (i) due to the original field, (ii) due to the induced current, are shown. The last chapter gives an account of modern work concerning the fascinating story of the atom; it has only been

touched upon briefly—just sufficient perhaps to whet a student's appetite for more, but not sufficient to distract him from the more fundamental parts of the subject.

As in the first edition, the treatment is mainly experimental and most of the graphs and numerical examples in the text are taken from actual observation. No attempt has been made, however, to give all the practical details of the experiments which students are expected to try for themselves, except in some of the more difficult exercises. In many instances graphical methods of dealing with experimental observations have been suggested.

In the near future, the author hopes to publish a text-book of practical physics of intermediate standard, and also a collection of examples, both worked and to be worked.

In all parts numerous diagrams will be found. These generally are in the form of a 'section,' and it is hoped that these will help in an understanding of the text and be found suitable for reproduction when occasion arises. Most of the original drawings have been executed from very sketchy material by my brother, Brigadier L. G. Smith, O.B.E., and to him I wish to express my very best thanks. The author also wishes to thank D. Orson Wood, Esq., M.Sc., for the valuable suggestions which he has continued to give. Thanks are also due to numerous correspondents who have pointed out errors of omission as well as of commission; also to Professor Sir Charles V. Boys, F.R.S., Professor A. Ferguson, D.Sc., Professor L. F. Bates, D.Sc., Dr. J. H. Brinkworth, Dr. H. J. T. Ellingham, and J. Nicol, Esq., B.A., B.Sc., who have made suggestions with regard to the earlier editions or who have gladly given advice when consulted. Lastly, the author would like to express his appreciation of the continued help given by his wife (*née* H. F. Taylor), without whose assistance in all matters connected with this edition its publication would have been delayed still further.

April, 1947.

Acknowledgments

The author would like to express his indebtedness to the following Scientific Societies or Publishers for permission to make use of diagrams and the relevant text occurring in their various publications:

The Physical Society—Figs. 3·8, 4·23. 39·17.
The Institution of Chemical Engineers—Fig. 15·13.
The Department of Commerce, Bureau of Standards, Washington —Figs. 8·4, 39·10 and 39·11.
Messrs. Macmillan and Co., Ltd.—The notes on 'Alcoholometry' taken from the Dictionary of Applied Physics.
Messrs. Longmans, Green and Co., Ltd.—Fig. 12·7.
Messrs. Sir Isaac Pitman and Sons, Ltd.—Fig. 3·1.

To Messrs. G. Bell & Sons, Ltd., and Professor Andrade, F.R.S., to reproduce from the latter's book, *The Mechanism of Nature*, the paragraph indicated on p. 336.

To the following for the loan of blocks:
Messrs. Baird and Tatlock, Ltd.—Fig. 2·7.
Messrs. Becker and Co.—Fig. 41·3(*a*).

To Messrs. Shell Mex Co., Ltd., for details concerning Fig. 4·15.

To Messrs. Siemens Electric Lamps and Supplies, Limited, for permission to reproduce Fig. 43·7.

To Messrs. Negretti and Zambra for Figs. 7·6(*b*), 48·14 and 48·15, together with details concerning them.

To Messrs. the Hewittic Electric Company, Limited, for Fig. 23·17.

To the Senate of London University and E. S. Cornell, Esq., M.Sc., to extract notes from Mr. Cornell's thesis on the History of Infra-Red Radiation.

To the Controller of H.M. Stationery Office for the tables of Logarithms and Trigonometrical Ratios from Professor J. Perry's *Practical Mathematics*.

And to the following Examining Bodies for permission to use questions set at various examinations held by them:
The Senate of London University (L.).
The University of Bristol (B.).
The Joint Matriculation Board of the Universities of Manchester, Liverpool, Leeds, Sheffield, and Birmingham (N.).

[The letters L, B, and N are used in the examples to indicate the source of the question. Also, I denotes Intermediate, S.C. denotes School Certificate, etc.]

THE GREEK ALPHABET

Letters.		*English Equivalent.*	*Names.*
Α	α	ă or ā	alpha
Β	β	b	beta
Γ	γ	g	gamma
Δ	δ	d	delta
Ε	ε	ĕ	epsilon
Ζ	ζ	z	zeta
Η	η	ē	eta
Θ	θ	th	theta
Ι	ι	ĭ or ī	iota
Κ	κ (ϰ)	k	kappa
Λ	λ	l	lambda
Μ	μ	m	mu
Ν	ν	n	nu
Ξ	ξ	x	xi
Ο	ο	ŏ	omikron
Π	π	p	pi
Ρ	ϱ	r	rho
Σ	σ, or (final) ς	s	sigma
Τ	τ	t	tau
Υ	υ	ŭ or ū	upsilon
Φ	φ	ph	phi
Χ	χ	kh	khi
Ψ	ψ	ps	psi
Ω	ω	ō	omega

CONTENTS

PART I

FUNDAMENTAL MEASUREMENTS AND THE GENERAL PROPERTIES OF MATTER

PART II

HEAT

PART III

OPTICS

INTERMEDIATE PHYSICS

PART I

FUNDAMENTAL MEASUREMENTS AND THE GENERAL PROPERTIES OF MATTER

CHAPTER I

THE MEASUREMENT OF LENGTH, ANGLE, TIME AND MASS

Natural Science.—That branch of human knowledge in which the properties of the material world are examined and then discussed is called natural science. Throughout the ages there have always been those who have endeavoured to become better acquainted with the events around them, whilst, until recently, there has always been a majority who have been content to live in the midst of phenomena about which they knew little ; to them science made little or no appeal. They were disposed rather to regard all phenomena as simple and self-explanatory. At the present time, however, such a state of affairs can hardly be conceived, since the advent of wireless and the extensive use of electricity in daily life have made it almost essential for everyone to become acquainted with the elements of science. But even in the centuries which have gone there have always been those who were not satisfied with a cursory view of Nature, so that they sought to discern the nature of things by careful experimental study. The experimentalist is for ever probing the inner secrets of Nature and, in so doing, he becomes more cognizant of the majesty and mystery of the universe around him. It is very probable that a study of Nature was begun soon after the appearance of Man upon this planet, for it is very difficult to imagine even amongst a tribe of uncouth savages an entire lack of interest in the wonders confronting it. To these people, however, every manifestation of Nature's power was a thing of awe and fear, capable only of being changed by prayer and intercession to the gods and demons which were believed to have their habitations in the material things of this world. Gradually, however, succeeding generations, profiting

by the knowledge handed down to them from their ancestors, began to refer various effects to certain fixed causes. They learned to interpret the signs of the heavens, and put their frail barks to sea when they thought that a period of calm was likely to persist. They became acquainted with the footprints of various animals and knew the times when these animals would come for water. Traps were set, and with the flesh of the animals so caught these people were able to provide for the sustenance of their families; the skins of the animals provided them with raiment for their bodies and also enabled them to erect a cover to protect themselves from the fury of the storm. These ancient inhabitants of the earth were really becoming familiar with laws, for they were realizing that certain causes would inevitably be followed by certain effects.

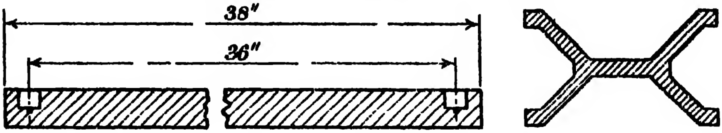
In every generation a few people were able to add a little to the sum of human knowledge, and each new discovery made further progress more rapid, until, during the last few decades, the advancement of scientific knowledge has been as remarkable as it has been beneficial to mankind. No man is able to claim a thorough acquaintance with all the laws and theories of modern science, so that it has been necessary to divide natural science into several branches, physics, chemistry, biology, etc., and the great strides which have been made in all these branches during the last century and this, have made further subdivision imperative in all these sections of natural science. In biology the properties of living matter are investigated. Here, much of the work is at present only of a qualitative nature, for the processes which are at work are very complicated and intricate, necessitating a vast amount of research before the laws governing them can all become known; recent developments in this field have made it very apparent that there are definite laws and that these laws must be obeyed or the penalty paid. In physics and chemistry the properties of inert matter are examined and the investigations now completed are so extensive that many quantitative laws are known, the discovery and formulation of which have been made possible by the availability of exact standards of measurement. These have arisen from the fact that, if real progress is to be made, exact comparisons must be possible. Amongst the instruments of greatest service in the development of modern science are the balance, thermometer, spectroscope, microscope, and that new and powerful tool the X-ray spectrometer—which has extended our knowledge of the structure of atoms in a manner which would otherwise have been almost impossible. Another reason why exact standards of reference have become so necessary at the present time is that industry is always making demands upon the scientist to supply it with

more accurate tools, or standards for checking the articles it manufactures. A mere mention of the aeroplane, or of the thermionic valve, brings home to us at once the truth of the above statements. We shall therefore begin our study of physics with a short discussion of the fundamental units which form the basis upon which modern science has been built. In passing, a brief reference to the difference between the science of to-day and that which flourished at the time of the ancient Greeks may not be inappropriate. Those early philosophers were content to make observations on a few things and proceed at once to develop a theory, and once having framed it they adhered to it most tenaciously. The procedure during the last few centuries has been very different. Scientists had realized that theories were utterly useless unless they could be substantiated by numerous facts, and they therefore set aside the art of making theories and directed all their attention to establishing facts. When these facts had been correlated, theories became possible, although modern scientists have always recognized that it is the facts which are true and that the theories are merely the product of man; hence, like man, they may be here to-day and gone to-morrow.

The Three Fundamental Units.—The statement that the height of St. Paul's Cathedral is 365 ft. conveys two ideas—one is the unit [the foot], while the other states how many times this unit is contained in the height of the object measured. Later on we shall find that all units, e.g. those of speed, force, electric current, pole strength, etc., may each be expressed in terms of three others, viz. length, mass, and time. These are the ***three fundamental units***, while all others are called ***derived units***. The fundamental units in scientific work are the centimetre, gram, and second, so that the system of units based on these particular units of length, mass, and time is referred to as the cm.-gm.-sec. [c.g.s.] system. In England and English-speaking countries, another system is nearly always used for domestic and commercial purposes: it is known as the foot-pound-second [f.p.s.] system because its fundamental units are the foot, pound, and second.

The Measure of Length.—In England the unit of length is the foot, which is defined as one-third the distance between the central traverse lines on two gold plugs in a bronze bar called the ***Imperial standard yard*** when this bar is at 62° F. and supported so that it is not bent when comparisons with it are being made. [Weights and Measures Act, 1878.] A longitudinal section of this bar is shown in Fig. 1·1 (*a*), the two gold plugs being shown in black. It will be observed that the upper surfaces of these plugs on which the fiducial lines are engraved are in the median plane of the bar where the errors due to any possible bending are a minimum.

The unit of length in the c.g.s. system is the centimetre which is defined as the one-hundredth part of the metre. This latter was intended to be one-ten-millionth part of the line of longitude passing through Paris and extending from the North Pole to the Equator. Actually this desire was not quite fulfilled, and so, for legal and scientific purposes, the metre is defined as the length at 0° C. between two fixed lines engraved upon the central flat portion of a platinum-iridium bar, a cross-section of which is indicated in



(a) (b)

FIG. 1·1.—Standards of Length.

Fig. 1·1 (b). This bar is termed the ***International prototype metre***. Its length in metres at any other temperature is given by

$$l_t = 1 + [(8\cdot651t + 0\cdot00100t^2) \times 10^{-6}]$$

where l_t is the length at $t°$ C.

The Measure of Mass.—In the British system the pound is the unit of mass; it is defined as the mass of a certain platinum cylinder marked 'P.S. 1844, 1 lb.', and deposited with the Warden of Standards in London. When a copy of this platinum standard is to be made in some other metal it is necessary to allow for the buoyancy of the air so that in recent acts the words 'in vacuo' have been added to define the standard condition of the platinum cylinder.

The metric system adopts as its standard of mass the gram, which is the thousandth part of a mass of platinum-iridium, called the ***International prototype kilogram***. This latter is very nearly the mass of a cubic decimetre of distilled water at such a temperature that its density is a maximum, viz. 3·98° C. when the pressure on the water is one atmosphere. It is just as necessary to specify the pressure as it is the temperature in the above statement, since the volume of a given mass of water depends upon the external pressure to which it is subjected.

Time.—The choice of a standard of time is more difficult than for the other fundamental units, since, whereas different lengths or masses may be compared with the same respective standard, no standard unit of time is available—time can only be measured by the repetition of a process. The rotation of the earth about its axis is an excellent standard of uniform motion, but it is not quite perfect. Tidal friction increases its period of revolution, while any contraction in its size tends to accelerate its motion. The

other natural clocks which astronomy offers to us are the revolutions of the planets or of the satellites of Jupiter. These are not convenient standards, however, so that they are only used as a last resort to confirm or disprove any variation which may have been suspected in some other standard clock.

The unit of time is the ***mean solar second*** which is the $\frac{1}{86400}$th part of a mean solar day. The solar day is the period which elapses between successive transits of the sun across the meridian at any point on the earth's surface. The duration of a solar day is not a constant magnitude but varies according to the time of the year when it is measured. It is for this reason that the average value of the solar day taken over a twelvemonth is used in defining our unit of time, and this mean value is called the ***mean solar day.*** Astronomers, however, use a different unit of time known as the ***mean sidereal second.*** This is derived from the mean sidereal day which is the average value of the period which elapses between successive transits of one of the fixed stars across a meridian, the average being taken over a period of one year.

In consequence of the earth's orbital motion round the sun, the time interval between two successive transits of the sun across the meridian at any place on the earth is different from that between two successive transits across that meridian of a fixed star. For simplicity, let us assume that the earth's orbit is a circle with the sun, S, Fig. 1·2, at its centre. This circle has a radius $9 \cdot 3 \times 10^6$ miles, and is described in 365 days, 6 hours, 9 min., 9 sec. (solar time)—the length of a so-called ***sidereal year,*** or the time interval between two successive appearances of the sun

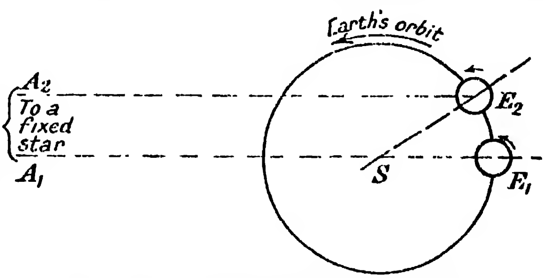


FIG. 1·2.—The Sidereal and Solar Days.

in the same position relative to the fixed stars. Let E_1 be the position of the earth when a transit of the sun and of a star occur simultaneously. When the earth has made one complete revolution about its axis, i.e. the next transit of the star takes place, it will be at E_2 but, as the diagram shows, a further rotation through $A_2\widehat{E_2}S$ must occur before the sun crosses the meridian. Hence the solar day is longer than the sidereal day. Actually the mean sidereal day is equal to 23 hours, 56 minutes, 4·09 seconds of mean solar time. [N.B.—If the earth rotated about its axis in the opposite direction, the solar day would be shorter than the sidereal day.]

The Vernier.—When it is desired to determine the distance between two given points it is quite fortuitous if that distance happens to be an exact multiple of the unit of length used; in general there will remain a fraction of a unit for which the relatively coarse divisions on the scale cannot account. This small fraction is determined with the aid of a vernier, the principle of which may be learned from the following: Let AB, Fig. 1·3 (*a*), be a line 9 cm. in length, and let AC and BD be two parallel lines each 10 cm. long. By dividing these two parallel lines into ten equal parts and joining corresponding points by straight lines the line AB is divided into ten equal divisions. This constitutes the vernier scale. Suppose now that the one extremity of a body being measured lies somewhere between the divisions marked 41 and 42 on the main scale. To locate the position of the end of the body more exactly the vernier scale is placed with its zero end in contact with the extremity of the object, when it is observed

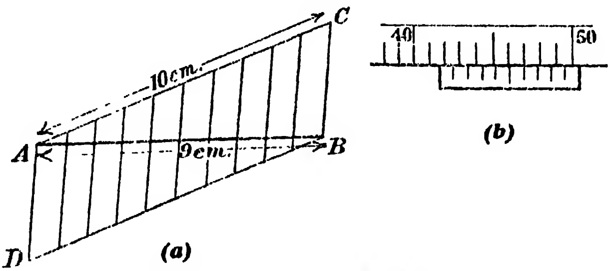


FIG. 1·3.—Principle of a Vernier.

that the sixth division on the vernier scale coincides with a division on the principal scale—cf. Fig. 1·3 (*b*). Since each division on the vernier is one-tenth of nine divisions on the principal scale, i.e. in this particular instance one-tenth of 9 cm., and therefore 0·9 cm., it follows that the difference between one division on the principal scale and one on the vernier is one-tenth of one division on the principal scale, i.e., 0·1 cm. on the vernier constructed above. Hence the difference between six scale divisions and six vernier divisions is 6 × 0·1 cm., so that the required reading is 41·6 cm.

In actual practice this method of constructing a vernier is always applied to the smallest divisions on the principal scale, i.e. one finds that the vernier is generally 9 mm. long, so that each division on it, if it is divided into tenths, is 0·9 mm. = 0·09 cm. The difference between one division on each of the two scales is then 0·01 cm. When greater accuracy is required, nineteen small divisions on the principal scale are divided into twenty parts so that the difference between one small division on it and one on the vernier is one-twentieth of a small division on the principal scale.

Slide Callipers.—As an actual example of the use of a vernier to determine tenths of a millimetre reference may be made to a pair of slide callipers, Fig. 1·4. Q is the main scale, graduated in cm. and mm., while the vernier V is attached to a movable jaw B. The jaws A and B are perpendicular to the scale Q; the body P whose length is required is inserted between these jaws. When the jaws are closed the zeros of the scale on Q and the vernier V should coincide, whilst when the jaws are open the position of the vernier zero gives, on Q, the perpendicular distance between the jaws. In this particular instance, 10 vernier divisions are equivalent to nine small scale divisions, i.e. to 9 mm., so that each vernier division is equal to 0·9 mm. Now the difference between one small division on the principal scale and one division on the vernier is [0·1 — 0·9 (0·1)] cm. = (0·1 — 0·09) cm. = 0·01 cm. From the figure it is seen that the diameter of the object P is between 7 mm. and 8 mm., and that the seventh division on the vernier coincides with a division on Q; hence the small fractional part which is required is the difference between 7 small scale divisions and 7 vernier divisions. This difference is seven times the difference between 1 principal scale division and 1 vernier division, viz. 7 × 0·01 cm. = 0·07 cm. The length of the object is, therefore, 0·7 + 0·07 = 0·77 cm.

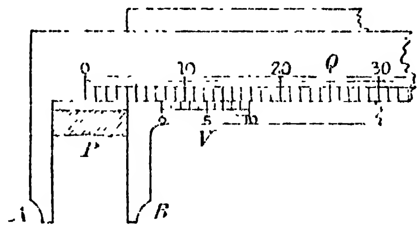


FIG. 1·4.—Slide Callipers.

The Micrometer Screw.—The micrometer screw gauge is another device for measuring small distances accurately. A linear scale in millimetres is engraved parallel to the axis of a cylindrical

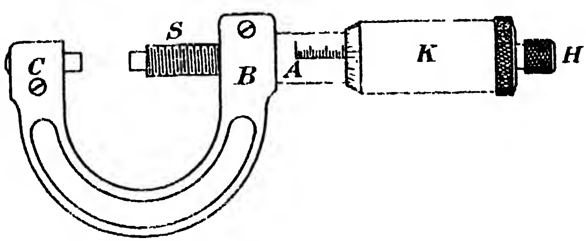
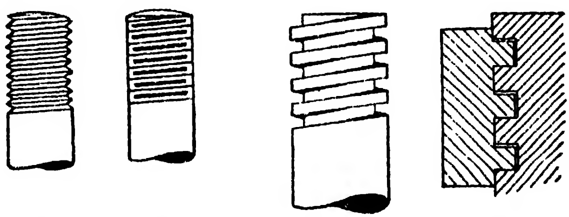


FIG. 1·5.—Micrometer Screw Gauge.

tube A, Fig. 1·5, this latter carrying a curved arm BC. Inside the tube A moves an accurate screw S, the pitch of which is 0·5 mm. —the pitch of a screw is equal to the length through which

the screw moves when it is rotated once about its axis. This movement is obtained by rotating the head H. This rotation also causes the collar K to turn around its own axis. The bevelled end of K is divided into 50 equal divisions, so that a rotation of K through one division corresponds to a movement of $\left(\frac{1}{50}\times 0{\cdot}5\right)$ mm. = 0·01 mm., these divisions being used for interpolating the distance between the mm. divisions on A. The extremity of S and the face of C are perpendicular to the axis of the screw; between these two jaws the object to be measured is placed. When these jaws are in contact the zero on the bevelled edge should coincide with the zero on the mm. scale; before using the instrument this point should always be tested, and if the instrument has a zero error the corresponding correction [1] must be applied. The head H is arranged so that when the jaws of C and S are in contact, either with each other or some object, a further rotation of H fails to impart any movement to K.

Screws.—The micrometer screw gauge just described is an example of the use which is often made of an accurately cut screw. The threads of screws are generally triangular or square in section as in Fig. 1·6 (*a*) and (*c*). Perhaps the most important thread used



(*a*) (*b*) (*c*) (*d*)

FIG. 1·6.—Screws.

by engineers is the ***Whitworth*** **V** thread in which the angle of the thread is 55 degrees—cf. Fig. 1·7. In all screws the distance through which the screw advances when it makes one complete revolution is called the ***pitch*** of the screw. In diagrams screws are conventionally represented as in Fig. 1·6 (*b*).

Back-lash.—Screws may be used to impart a translatory motion

[1] The words *correction* and *error* are sometimes used as if they were synonymous. This is not so, it being preferable to define the correction as the quantity which must be *added* algebraically to the observed reading in order to obtain the true reading. The error is then equal to the negative value of the correction. Thus if a thermometer reads − 0·6° C. when in melting ice, the temperature of which is defined as 0° C., then the correction is + 0·6° C. and the error − 0·6° C.

to a nut in which they work, and the amount of rotary motion which can be imparted to a screw without causing any movement of the nut is known as back-lash. It is sometimes due to wear or to imperfections in the manufacture. Very often, however, especially if the screw is intended for precision work, a certain amount of back-lash is allowed when the screw is being cut. The reason for this is that if an attempt is made to make the screw and nut fit exactly then the fit is soon destroyed by wear owing to the somewhat large forces operating upon screw threads.

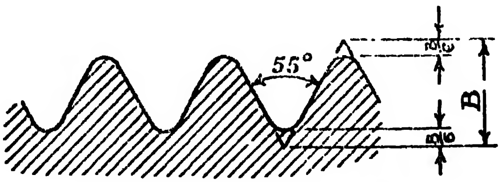


FIG. 1·7.—Whitworth Screw Thread.

It is therefore better to design the screw so that only one of its faces is in contact with the nut—in this way the wear is reduced to a minimum. When using screws for the purpose of estimating small distances, care must always be taken to turn the screw in one direction through a relatively large distance when setting the screw before an observation. Fig. 1.6 (*d*) will perhaps help to make these remarks more clear.

The Travelling or Vernier Microscope.—In order to measure short vertical or horizontal distances, a vernier microscope is frequently used. A precision form of this instrument is shown in Fig. 1·8. It consists of a microscope, M, clamped to a tube, A, supported in a rigid frame, B. When the microscope is thus clamped, a maximum displacement of 4 cm. may be imparted to it by means of a screw, S, operated by the milled head, H. The amount of this displacement is measured by a horizontal scale, D, which gives the complete number of revolutions of the milled head, the fractional part of a rotation being given by the divisions on the wheel, F, attached to the screw. The microscope may, however, be clamped in position at any point along the tube so that longer distances are measured by a succession of smaller displacements of the microscope. The microscope is fitted with an achromatic objective and eye-piece with cross-wires. The object under examination is supported on a small sliding table, resting upon geometric clamps, and provided with aligning adjustments operated by screws. A steel spring placed inside the tube, A, and attached to the stud, K (fixed to the stand, B), and to the end, L, of the tube, so that the spring is stretched, keeps the end, N, of the tube, A, in

contact with the extremity of the screw, S. [When the instrument is to be used to measure vertical distances, it is provided with a tripod base with levelling screws, so that the microscope may be traversed vertically.]

In order to measure with the aid of this instrument the diameter of a brass disc, for example, the eye-piece is first adjusted so that the cross-wires can be seen clearly—when this adjustment has been performed the eye-piece must not be disturbed again. The microscope is then raised or lowered by the rack and pinion, R, until the details in the object are visible. This having been done, the microscope is moved either horizontally or vertically until the image of the edge of the disc coincides with the cross-wire in the eye-piece, the microscope being adjusted so that there is no parallax. The appropriate scale is then read, and the microscope

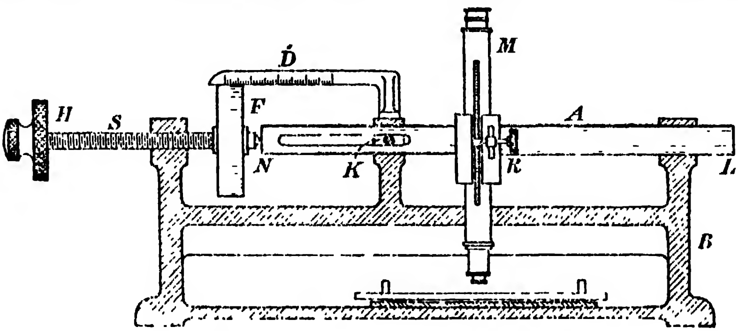


Fig. 1·8.—Travelling or Vernier Microscope.

afterwards moved so that the other extremity of the object coincides apparently with the cross-wire. The difference in the readings on the scale gives the diameter required.

Whenever it is necessary to determine the diameter of a circular object it is always advisable to measure two diameters in directions at right angles to each other. If the body has a cross-section which is slightly elliptical its mean diameter is the mean of any two diameters at right angles to one another, for it may be shown that the mean of any two mutually perpendicular diameters of an ellipse which is almost a circle is a constant for any given such ellipse.

Experiment.—The following exercise, which is to determine the number of centimetres equivalent to one inch, provides a means of becoming familiar with the use of a travelling microscope and illustrates also how to obtain a mean value from a series of observations. A steel scale graduated in inches and tenths of an inch is attached to the bed of the microscope and the microscope focussed on an image of one of the dividing lines on the inch scale, the eyepiece having previously been adjusted so that the cross-wires in the microscope are

clearly seen. The microscope is arranged so that there is no parallax between the cross-wires and the image of the particular dividing line which is being observed. The reading on the centimetre scale attached to the microscope is noted. The next dividing line is then observed and a similar reading obtained. The process is continued until about ten observations have been obtained. The fractional part of a centimetre corresponding to one-tenth of an inch may be found as follows. Let us suppose that when the microscope is moved by successive increments (tenths of an inch), the corresponding readings on the scale of the travelling microscope are $a_1, a_2, a_3, \ldots a_{10}$ (say). How are we to obtain arithmetically the best value for the shift of the microscope corresponding to 0·1 inch? If we deduce $(a_2 - a_1)$, $(a_3 - a_2)$, etc., and then calculate the mean of these quantities we only utilize the first and last observations, for

$$(a_2 - a_1) + (a_3 - a_2) + \ldots + (a_{10} - a_9) = (a_{10} - a_1).$$

This may be avoided by calculating

$$(a_6 - a_1), (a_7 - a_2), \ldots (a_{10} - a_5).$$

The mean of these quantities, viz.,

$$\tfrac{1}{5}[(a_6 + a_7 + \ldots a_{10}) - (a_1 + a_2 + \ldots a_5)]$$

then gives the average length in centimetres equivalent to 0·5 inch, and we notice that the calculation involves each reading once, and once only.

The Spherometer.—This instrument, which was specially designed for determining the radii of curvature of spherical surfaces, is shown in Fig. 1·9 (*a*). The circumference of the screw head H is divided into 100 equal divisions; the pitch of the screw is 0·5 mm., so that each division corresponds to 0·005 mm. The scale S gives the number of complete revolutions which the head makes. The points of the legs A, B, C are all in one plane and form an equilateral triangle. To set the instrument when, for example, the radius of curvature of a convex surface is being determined, it is placed on a sheet of glass and the screw-head, H, turned until D, the central leg of the instrument, is a little lower than A, B and C, Fig. 1·9 (*b*); if the instrument is tapped gently it rocks about an axis passing through the points of contact with the surface of D and of one of the fixed legs. The screw-head N is moved until this rocking ceases—the points A, B, C and D are then all in one plane. In estimating the position where the rocking just ceases, rotate the head through five divisions at a time when the position is being approached. Do this until the rocking ceases. Then, having rotated the head back through several divisions to avoid errors due to back-lash when the screw is advanced, turn the head in the initial direction through one division. Test for rocking on each occasion and proceed until the exact position of no rocking is located between two successive readings of the position of the head H. H is then raised until, when the spherometer is placed on the spherical surface, the instrument just ceases to rock. The

difference in the readings gives the height through which the screw D has been moved—let this be h. Let a be the distance between the outer legs and D when all four legs have their extremities in one plane. Then in Fig. 1·8 (c)

$$OA^2 = AD^2 + OD^2$$

i.e. $$R^2 = a^2 + (R - h)^2 = R^2 - 2Rh + h^2 + a^2$$

i.e. $$R = \frac{a^2}{2h} + \frac{h}{2}.$$

If s is the distance AC, $s = a\sqrt{3}$, so that

$$R = \frac{s^2}{6h} + \frac{h}{2}.$$

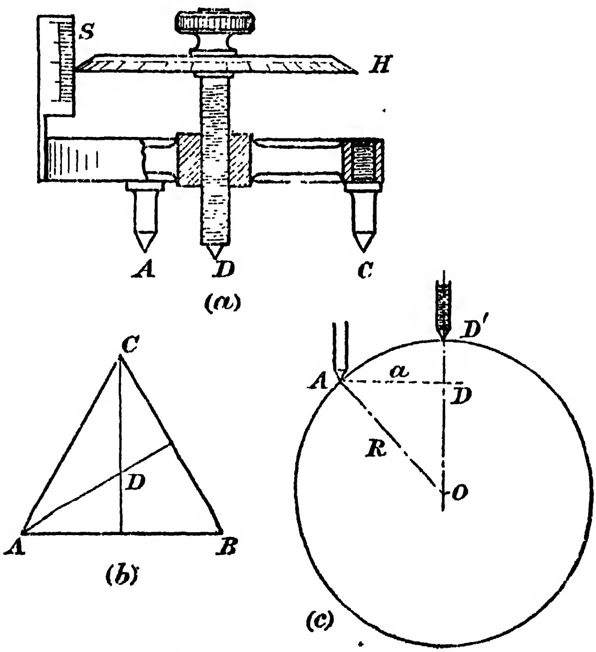


FIG. 1·9.—The Spherometer.

In using this instrument it is convenient to obtain first a reading with the central leg in the lowest position which it will occupy in any experiment, for, unless this procedure is adopted, the fractional parts are (1 — the reading on H); this tends to be confusing. The head is then rotated so that the screw D moves upward. The number of complete revolutions made by this head is best determined by counting, since it is not always easy to locate the position of the periphery of H on the vertical scale S. The fractional part of a rotation is determined in the usual way.

Example.—In determining the thickness of a piece of glass with a spherometer the pitch of whose screw is 0·5 mm., the readings were (i) 0·47 mm., (ii) three complete turns and a fractional part 0·25 mm.

$\therefore$ Thickness $= [(3 \times 0{\cdot}5) + 0{\cdot}25] - 0{\cdot}47$ mm. $= 1{\cdot}28$ mm.

Circular Verniers.—When it is essential to measure angles with an accuracy greater than that obtainable with a protractor, use is made of a circular vernier. The actual vernier found on any particular instrument will depend upon the smallness of the divisions on the main scale. As an example let us assume that this scale reads directly to half a degree. If the ultimate aim is to measure an angle correct to one minute the following procedure may be adopted. Twenty-nine divisions on the main scale are divided into thirty equal small divisions, so that the difference between one scale and one vernier division is one-thirtieth of half a degree, i.e. one minute. Hence if the 12th division on the vernier coincides with a division on the main scale, the fraction of a degree to be added to the reading of the main scale is 12 minutes. [Note that the main scale is divided into half-degree divisions.]

Indirect Methods of Measuring very great Distances.—Hitherto only direct methods of measuring a length, i.e. methods involving the repeated application of a standard or sub-standard rod of known length, have been mentioned. Let us see whether it is legitimate to apply indirect methods to measure lengths, such as the distance between two mountain peaks or that between the earth and a heavenly body. In these indirect methods a base line is selected—it may be the diameter of the earth's orbit—and various angles are measured. The required distance is then calculated by means of some trigonometrical formula. In this method certain light rays are identified with straight lines defining the sides of a triangle: but this is an assumption, finally to be tested experimentally. It is now known that for terrestrial distances the assumption is justified, but in an astronomical survey involving immense distances certain corrections have to be applied. These remarks are made here in order to show that there may be inherent difficulties when indirect methods are used to determine an apparently simple physical quantity such as a length.

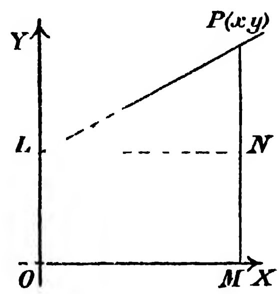


FIG. 1·10.

The equation $y = a + bx$.—Let PL, Fig. 1·10, be a straight line making an intercept of length a on the axis OY. Let P be any

point $(x, y,)$ on this line. Draw PM perpendicular to OX and LN parallel to OX. Then

$$y = \text{PM} = \text{PN} + \text{NM}$$
$$= \text{PN} + \text{OL}$$
$$= \frac{\text{PN}}{\text{LN}} \cdot \text{LN} + \text{OL}$$

If we call $\frac{\text{PN}}{\text{LN}} = b$, this equation becomes

$$y = a + bx.$$

Any equation of this type therefore represents a straight line, i.e. it is a ***linear equation*** between the variables x and y; if an equation contains powers of x it may be said at once that it is the equation to some form of curve. The constant b in the above equation measures the ***slope*** of the line.

The Graphical Determination of Laws.—Whenever possible the results of an experiment should be shown graphically and if the results happen to lie on a straight line its equation determined. The constants in such an equation will often convey useful information to us. Moreover, if all except one or two of the points lie on a straight line, then such a graph tells us what observations should be repeated, for they are most likely to be in error. If the points do not lie on such a line, a fact which is best revealed by stretching a piece of black cotton across the paper, it may be advantageous to plot the logarithms of one or both of the quantities involved. The method of attack in such an instance may be gathered from the following example.

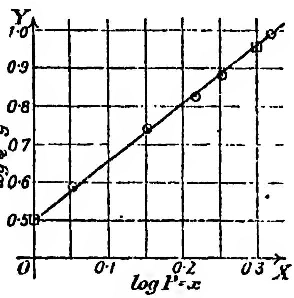


FIG. 1·11.

Example. The following numbers were obtained in a certain laboratory experiment.

Q	3·84	5·43	6·80	7·56	9·82
P	1·13	1·42	1·66	1·80	2·11

Discover the law connecting Q and P.

The graph obtained by plotting Q, as ordinate, against P is not a straight line. If, however, log Q is plotted against log P, as in Fig. 1·11, it is apparent that these two quantities are related to each other by a linear law. The intercept on OY is 0·50, while the slope is 1·51. [In measuring the slope of a line always choose two points on the line as far apart as possible—shown ⊡.] If we call $\log P = x$, and $\log Q = y$, the equation to the line is $y = 0{\cdot}50 + 1{\cdot}51x$, i.e. $\log Q = 0{\cdot}50 + 1{\cdot}51 \log P$.

The relationship between Q and P is therefore

$$Q = 10^{0\cdot50}P^{1\cdot51}$$

or

$$Q = 3\cdot18P^{1\cdot51}.$$

The Measurement of Mass.—The mass of a body is usually found by comparing that body with a set of standard masses, invariably referred to as a 'box of weights.' The comparison is carried out by means of a balance, which is really an equi-arm lever poised about a fulcrum. The masses to be compared are placed in pans which are suspended on knife-edges from the extremities of the beam. The accuracy of the balance depends, to a large extent, upon the design of the beam, which must be light but rigid. It must be light in mass if the sensitivity of the balance is to be high, and yet sufficiently rigid that its shape is not deformed under the greatest load for which the balance has been designed. The knife-edges are usually made of agate, this substance being chosen on account of the facts that it is hard, does not tarnish, and may be worked until a straight edge has been obtained.

The equality of the masses is ascertained by observing the deflections of a long vertical pointer perpendicular to the beam. When this pointer swings through the same distance on either side of its zero position, then the two masses in the pans are equal. The balance is protected in a glass case and the humidity of the atmosphere inside the case is greatly minimized by the use of concentrated sulphuric acid or solid calcium chloride contained in a glass receptacle. When the balance is not in use the beam and pans are not free to move, the beam being raised so that there is no permanent load on the knife-edges, whilst the pans rest on supports, all these conditions being obtained by the rotation of a small handle or wheel, which is outside the balance case.

Measurement of Time.—The evolution of watches and clocks is a result of man's desire to divide the mean solar day into smaller intervals of time. Most clocks depend upon the motion of a pendulum which, in its most simple form, consists of a heavy bob fixed to the end of a string, the other end being attached to some definite point. If the size of the bob is small compared with the length, l (cm.), of the string, then the time, T (sec.), of a complete swing, i.e. the time which elapses between the successive transits of the bob in the same direction past some fixed point or line, is given by the equation [cf. p. 38]

$$T = 2\pi\sqrt{\frac{l}{g}},$$

where g is the intensity of gravity [981 cm. sec.$^{-2}$ in England]. This equation enables the length of a seconds pendulum to be found,

this being the special pendulum making half a complete swing each second so that $T = 2$ sec. Its length is given by

$$l = \frac{T^2 g}{4\pi^2} = 99{\cdot}3 \text{ cm.}$$

Errors of Observation.—In this introductory chapter some remarks should be made concerning the accuracy of one's experimental results. The instruments available for determining the quantity in question should be critically examined to see what accuracy they can give, and care must be exercised to see that the final result recorded does not express an accuracy beyond the limits which the instruments can give. Suppose, for example, that the dimensions of a rectangular piece of wood are found by one student to be 3·96 cm. × 4·72 cm. × 1·74 cm. He may state that the volume is 32·522688 cm³. Is this result justifiable? Let us suppose that a second student measures the same block with the same pair of callipers. His observations are 3·98 cm., 4·75 cm., and 1·71 cm. respectively. If he proceeds to calculate the volume as the first student did, i.e. without thinking what he is doing, he will obtain 32·327550 cm.³ Why the difference? It is simply because their observations, just like all other observations, are subject to error so that they were not justified in stating anything more than 32·5 cm.³ and 32·3 cm.³ respectively.

It must be emphasized that when the first student asserts that the volume is 32·5 cm.³, he implies that his result is accurate to within 0·1 cm.³, i.e. the error is ± 0·1 cm.³ If, for example, there were reason to suppose that the error were ± 0·4 cm.³, he should indicate it by writing the result as (32·5 ± 0·4) cm.³

These few remarks may be further exemplified by extracts from the writings of a correspondent to *The Times*, 22 February, 1928. A few days before the result of a speed record had been quoted as 206·95602 m.p.h., a figure which implies that the speed was more than 206·95601 and less than 206·95603 m.p.h. If it does not, then the last figure has no meaning. To use this figure means that the error in the observations did not exceed one part in 20,695,602; thus, if the speed were measured over a distance of one mile exactly, then the timing gear gave results accurate to within one-millionth of a second. If, on the other hand, the timing gear was absolutely accurate, the distance of one mile must have been accurate to one three-hundredth of an inch. The correspondent ends by saying that a round figure of 207 m.p.h. is about the utmost that the actual measurements are likely to justify.

Errors due to Parallax.—The accuracy of the observations, for example, of the positions of the two points, whose distance apart is required, made with an ordinary scale graduated in cm. and

mm. is limited by the fact that the graduation marks have a finite thickness and because the eye cannot estimate accurately differences of length less than 0·1 mm. The error introduced in the measurement of a length in this way is likely to be at least 0·2 mm. Frequently, however, the observations may be much less accurate unless precautions have been taken to eliminate 'errors due to parallax'—i.e. to the apparent change in the position of an object with reference to some fixed object due to a change in the position of the observer. Such errors arise in the present instance if the scale is used as in Fig. 1·12 (*a*) owing to the thickness of the bar on which the scale is engraved. Thus, with the eye at E_1, the position of A appears to be 9·1 cm.; from E_2 it is 9·2 cm., etc. To eliminate such errors the graduated edge of the scale should be placed in contact with the points between which the distance is to be measured—cf. Fig. 1·12 (*b*).

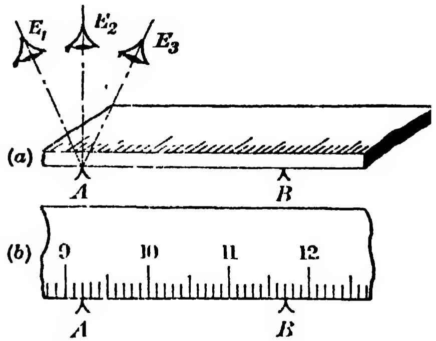


FIG. 1·12.—Errors due to Parallax.

EXAMPLES I

1.—Calculate the circular measure of an angle of 47°. By means of a diagram calculate its sine, cosine, and tangent. What is the angle whose circular measure is unity?

2.—Describe a vernier and its use on a pair of callipers.

3.—Describe, with the aid of a diagram, a micrometer screw gauge.

4.—If the diameter of the earth were increased by 1 ft., calculate the increase in its circumference.

5.—Calculate the length of a simple pendulum which will beat half-seconds.

6.—Criticize the following: A rectangular block measures 7·16 cm. × 6·73 cm. × 4·05 cm. Its volume is therefore 195·1565 cm.^{3}

7.—Discuss the advantages of representing a series of observations by means of a graph. How would you plot a series of observations of the time taken by a trolley to travel different distances down an inclined plane, so as to bring out the law involved as clearly as possible?

8.—Describe a spherometer and deduce the formula necessary when using this instrument to determine the radius of curvature of a spherical surface.

CHAPTER II

THE ELEMENTS OF DYNAMICS

Mechanics.—The science of mechanics deals with the properties of bodies, and it is usually studied under the headings, or sections, called ***Dynamics, Statics*** and ***Hydromechanics.*** The first branch, viz. ***Dynamics,*** deals with bodies which are in motion relative to their surroundings ; the second, viz. ***Statics,*** concerns bodies at rest, whilst ***Hydromechanics*** is the science of liquids. This latter subject is again subdivided into ***Hydrostatics*** and ***Hydrodynamics,*** the former section dealing with liquids at rest whilst in ***Hydrodynamics*** the motion of liquids is studied.

The Material Particle.—At this stage, perhaps, reference ought to be made to the size of the objects we discuss in an elementary treatment of the subject of dynamics. The equations which are deduced only apply to a body which is so small that it may be regarded as a mathematical point—such a body is known as a ***material particle.*** If the body is not small, attention must be paid to the fact that it is capable of rotation and therefore possesses energy due to rotation. A material particle may therefore be defined as a body which is so small that its energy of rotation may always be neglected.

Motion in a Straight Line

Displacement.—The motion of a body is only detected by observing its position with reference to its surroundings. If the position of the body changes with reference to its surroundings, then that body is said to be undergoing a ***displacement.*** Thus, if at some initial time, a body—here represented by a point—is in the position O, Fig. 2·1(*a*), and later on it is at A, then the body has been displaced, and the

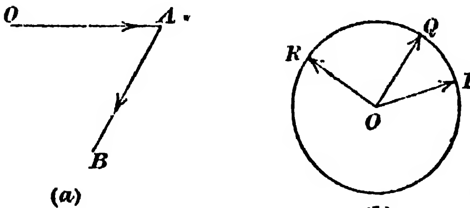


Fig. 2·1.

displacement is expressed by the length and direction of OA. At some later period in its history the body may be at B, so that its further displacement is AB, whilst the actual displacement from the origin is OB. The idea of displacement must always convey that of direction as well as that of magnitude. Thus, if O, Fig. 2·1 (*b*), is the original position of a body, whilst P, Q, R, etc., points on the circumference of a circle whose centre is O, are its subsequent positions, then the magnitude of the displacement is fixed but the direction is variable.

Sense of Direction.—Every displacement has magnitude and direction; they all have 'sense' too; for example, a body may be displaced from O to A, or from A to O. The sense of the direction is opposite in these two instances, the sense of the displacement being indicated in a diagram by the use of small arrow-heads.

Representation of Displacement.—A displacement is represented on a drawing by a straight line whose direction and sense are that of the displacement and whose length is proportional to the magnitude of the displacement. Any quantity which can be represented in magnitude, direction, and sense, is called a ***vector***. Other quantities are ***scalars***. Velocity, force, magnetic intensity, etc., are vectors, while potential, energy, money, etc., are scalars.

Relative Displacement and Rest.—If two trains are moving in the same sense along parallel tracks, a passenger in one of the trains may observe the following facts:—If he is travelling in the faster train, the other train will appear to him as if it were receding, whilst if he is in the train which is moving less rapidly, he will observe a forward motion of the faster train. ***Relative displacement*** is defined as the displacement of one body with respect to another. In fact all displacements must of necessity be relative ones, although we are accustomed to think of absolute displacement because our idea of rest is generally associated with non-moving bodies on the surface of the earth. Actually the earth is moving on its own axis, round the sun and through space, so that when it is said that a body is at rest, the statement is only intended to convey the fact that its displacement, with respect to its surroundings [generally on the surface of the earth], is zero.

The Composition of Displacements.—In order to fix our ideas let us consider the displacement of a marble which rolls across the floor of a moving carriage. Let PQRS, Fig. 2·2, be the carriage, while O is the initial position of the marble. Let us further assume that in the time required for the marble to travel from O to B the point in the carriage corresponding to O has moved to A, i.e.

A occupies the same position relative to the final position P′Q′R′S′ of the carriage as O did with respect to the initial position PQRS. If C is the point corresponding to B, it follows that C is the final position of the particle and that the actual displacement is completely represented by OC. We say that it represents the actual displacement, but it must be understood that this implies the existence of a reference body absolutely at rest. Such a body is unknown. The result just obtained is a particular instance of a general theorem—we refer to the ***parallelogram law of vectors*** which may be stated as follows:—***Two vectors of the same type may be added together by constructing a parallelogram***

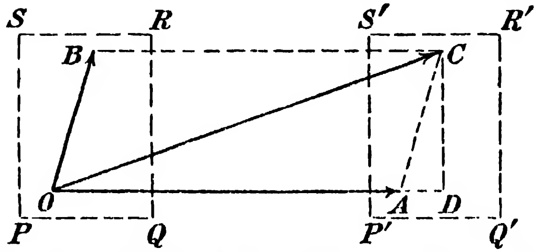


FIG. 2·2.—The Composition of Displacements (Vectors).

the adjacent sides of which are proportional to the vectors. The diagonal drawn through the point of intersection of these two sides represents the resultant of these two vectors completely. Thus OC, Fig. 2·2, is the resultant of the two like vectors OA and OB. The vectors OA and OB are said to have been added vectorially.

The magnitude of the resultant is easily found, for if CD is drawn perpendicular to OA to meet OA produced in D, then

$$\begin{aligned} OC^2 &= OD^2 + CD^2 \qquad [\because \widehat{ODC} = 90°] \\ &= (OA + AD)^2 + CD^2 \\ &= OA^2 + (AD^2 + CD^2) + 2 \,.\, OA \,.\, AD \\ &= OA^2 + AC^2 + 2 \,.\, OA \,.\, AC \cos\theta \\ &= OA^2 + OB^2 + 2 \,.\, OA \,.\, OB \,.\cos\theta, \end{aligned}$$

where θ is the $\widehat{AOB}$.

Speed.—The idea of speed is obtained by associating the conception of time with that of displacement. If, for example, a ship goes from one port to another in one day, while another occupies two days for the same journey, then the speed of the former vessel is twice that of the second. The speeds which are referred to here are average values of the speeds of the vessels, because the vessel starts from rest and comes to rest at some other point, so that its actual speed at some times will have been less than its average

speed, whilst at others it will have been greater. The average speed of a body is given by the expression

$$\text{average speed} = \frac{\text{distance traversed}}{\text{time occupied in so doing}}.$$

[Care must be taken to avoid such expressions as 'the speed of the ship was 22 knots per hour,' since ***knot*** is a nautical term used to imply a speed of one sea-mile per hour. A sea-mile is intended to be such a distance on the earth's surface that an angle of one minute is subtended by that arc at the earth's centre. The British Admiralty takes this to be 6020 feet.]

Velocity.—When a body moves in a definite direction the speed of the body in that direction is called its ***velocity.*** It is important to remember that velocity always implies speed in a fixed direction.

Uniform Velocity. The term ***uniform velocity*** is used to convey the idea that the distance traversed in any small interval of time is the same for all such intervals, however small the interval of time may be. Thus, if a body moves in a given direction 20 ft. in 5 secs. it is not justifiable to say that its velocity is uniform, for it is conceivable that if the position of the moving object had been observed at the end of every second, say, then the displacements in those seconds may have been found, for example, to be 3, 5, 6, 4 and 2 ft. respectively. The velocity (strictly, the mean velocity) in the first second is 3 ft. sec.$^{-1}$; in the second second 5 ft. sec.$^{-1}$; in the third second it is 6 ft. sec.$^{-1}$, etc., whilst the average velocity over the interval is 4 ft. sec.$^{-1}$ But if the velocity had been uniform and the body had been displaced 20 ft. in 5 secs., then in 1 sec. the displacement would have been 4 ft.; in 0·5 sec. 2 ft.; in 0·25 sec. 1 ft., etc.

Velocity-time Curve or Graph.—Suppose that a particle has an initial velocity of 2 ft. sec.$^{-1}$ and that at the ends of the first 3 sec. of its motion its velocity is 3·4, 4·0 and 4·2 ft. sec.$^{-1}$ respectively. Its motion can be shown graphically as follows: axes OX and OY are chosen at right angles to each other; one division along OX representing 1 sec., while one division along OY represents a velocity of 1 ft. sec.$^{-1}$—cf. Fig. 2·3 (*a*). The point A indicates the initial velocity, while the points B, C and D indicate the subsequent velocities of the body. Now the points A, B, C and D can be joined together either by short straight lines [shown dotted] or, since it is legitimate to presume that the velocity of the body does not change abruptly but continuously, they may be joined together by means of a smooth curve. This smooth curve is called the ***velocity-time curve.***

It is now necessary for the significance of the area under this

curve, i.e. the area OABCDRQP, to be found and its meaning interpreted. Since velocity means displacement per unit time, the statement that the uniform velocity of a body is 10 ft. sec.$^{-1}$ implies that the distances traversed in 1, 2, 3 and 4 secs. of its motion will be 10, 20, 30 and 40 ft. respectively. In Fig. 2·3 (*b*) the velocity-time curve for such a motion has been constructed; this curve will be the straight line AE, because the velocity is constant. Now an area of 1 sq. unit represents a distance traversed of 5 ft., because the unit length parallel to OX represents 1 sec. and

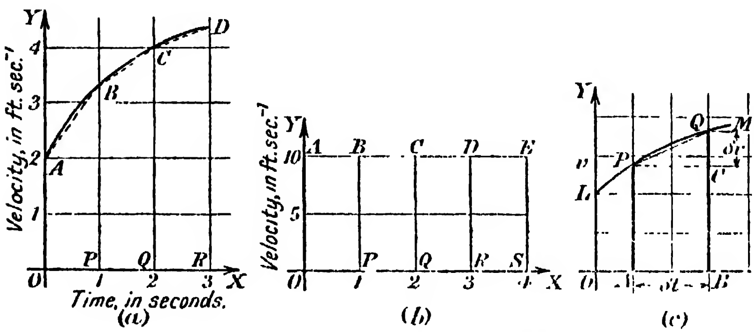


FIG. 2·3.—Velocity-time Curves.

the unit length parallel to OY represents a velocity of 5 ft. sec.$^{-1}$, so that the area OABP indicates a distance traversed of 10 ft., since OABP is 2 sq. units in area. The complete rectangle has an area of 8 sq. units, so that the distance traversed in 4 sec. is $8 \times 5 = 40$ ft., a value which agrees with that obtained from the definition of velocity [cf. p. 21].

Now when the area under the curve is an irregular one, as in Fig. 2·3 (*a*), it is still true that this area represents the distance through which the body has moved. In this particular example the area is 10·51 sq. units, and since 1 sq. unit represents a distance traversed equal to 1 ft., the actual distance traversed is 10·51 ft.

Algebraic Formula for Distance Traversed.—If the velocity of a body is uniform and equal to u ft. sec.$^{-1}$, the velocity-time curve will be a straight line parallel to the time-axis. If unit distance along the axis OX represents 1 sec. and unit distance along OY represents unit velocity, after t sec. the area of the velocity-time curve will be ut units of area, so that s, the distance traversed is ut, because unit area represents unit distance,

i.e. $$s = ut.$$

Acceleration and Retardation.—Consider the velocity-time curve obtained from the following observations:—

Time in seconds . . .	0	1	2	3	4	5
Velocity in ft. sec.$^{-1}$ . .	4·0	4·5	5·0	5·5	6·0	6·5

The graph is shown in Fig. 2·4, and it is at once apparent that the 'curve' is a straight line, i.e. it is linear. Whenever the velocity-time graph is linear, the motion is said to have been ***uniformly accelerated*** if the velocity has been increasing, and ***uniformly retarded*** if the velocity has been diminishing. These terms indicate that the velocity has been increased or diminished by equal amounts in equal intervals of time.

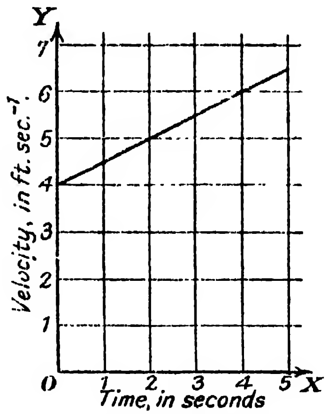


FIG. 2·4.—Velocity-time Curve for Uniformly Accelerated Motion.

Definition.—***Acceleration, i.e. the rate of change of velocity, is said to be uniform when the velocity increases by equal amounts in equal intervals of time, however small those intervals may be.***

A glance at the above table shows that the change in velocity is 0·5 ft. per sec. every second; or, as it is more usually written 0·5 ft. per sec. per sec., or 0·5 ft. sec.$^{-2}$.

When the velocity-time curve is not a straight line the velocity does not increase by equal amounts in equal intervals of time; i.e. the acceleration is ***non-uniform***. The acceleration at any particular instant may be determined as follows. Let P and Q, Fig. 2·3 (c), be two neighbouring points on a velocity-time curve LM. Draw PA and QB parallel to OY, and PC parallel to OX. Then PC and CQ represent small increments in the time and velocity respectively: we denote them by δt and δv. Since these two quantities are small their ratio $\dfrac{\delta v}{\delta t}$ is the mean acceleration of the moving particle during the interval of time when its motion is represented by P and Q. As the points P and Q move toward one another, the ratio $\dfrac{\delta v}{\delta t}$ remains finite and approaches a limiting value which is the acceleration of the particle at P. This limit is denoted by $\dfrac{dv}{dt}$, or $\dot{v}$, and measures the slope of the tangent to the curve at P. Thus, to determine the acceleration at a given time, the tangent

at the corresponding point on the velocity-time curve must be drawn. Its slope gives the acceleration required. In general, it is difficult to draw a tangent accurately, so that to obtain the acceleration at any instant we must measure the slope [acceleration] of the curve at several points and construct a curve showing the relation between the slope and the time. When this curve has been drawn the acceleration at any instant may be read off at once and the value obtained will be more reliable than that found by drawing the tangent at the point in question, since several tangents have been drawn in constructing the final curve.

Angular Velocity and Angular Acceleration.—When a particle, P, is rotating in a plane about a fixed point O the rate at which the angle between OP and a fixed line OX changes is termed the ***angular velocity*** of the point P. It is generally denoted by the symbol ω. The rate of change of the angular velocity is called the ***angular acceleration***, (α), i.e. $\alpha = \omega$.

Algebraic Formulæ for Uniformly Accelerated Motion

To find the Velocity after a Given Time.—Let u be the initial velocity, a the acceleration, t the time and v the final velocity of a moving body, all these quantities being expressed in the same system of units. Now a, the acceleration of the moving body, is the increase in velocity per unit time, so that the velocity at the end of the first unit time interval is $u + a$. Similarly the velocity at the end of the second unit interval will be $(u + a) + a$, or $u + 2a$. Hence, at the end of time, t, the velocity will have increased by an amount at, so that the actual velocity which has already been called v is then $u + at$.

$$\therefore v = u + at.$$

To find the Space Traversed in a Given Time Interval.—If the motion of the particle is represented by the symbols used in the last paragraph, the velocity-time graph is easily constructed; since the acceleration is uniform, we know from the definition of such an acceleration that the graph must be a straight line. This is represented by AB in Fig. 2·5. If BC is drawn perpendicular to the axis OX the area OABC represents the space traversed in time OC = t.

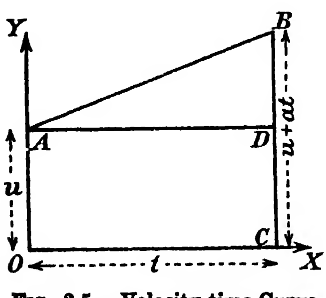


Fig. 2·5.—Velocity-time Curve for Uniformly Accelerated Motion.

$$\begin{aligned}\text{Now area OABC} &= \text{rect. OADC} + \Delta\ \text{ABD} \\ &= \text{OC . CD} + \tfrac{1}{2}\text{AD . DB} \\ &= ut + \tfrac{1}{2}at^2\end{aligned}$$

and this is s; consequently $s = ut + \frac{1}{2}at^2$.

The two equations which we have proved can be combined to form a new one. We have

$$\begin{aligned}v^2 &= (u + at)^2 \\ &= u^2 + 2a(ut + \tfrac{1}{2}at^2) \\ &= u^2 + 2as.\end{aligned}$$

Motion under Gravity.—When a body moves either towards or from the surface of the earth it is said to be moving under the influence of gravity. Such motions have been considered from the days of the Greeks, foremost among whom was Aristotle, who taught that heavier bodies fell towards the earth more rapidly than lighter ones. The validity of this doctrine was not disputed until Galileo [1564–1642] showed that two bodies, the mass of one being ten times that of the other, fell together when released at the same instant. These experiments, which were conducted from the leaning tower of Pisa before the eyes of his opponents, may be said to have sounded the 'last post' over the old doctrines which were founded on speculation, and which, in time, have been superseded by teachings based upon experimental fact. The ideas of Galileo were developed by Sir Isaac Newton [1642–1727], who devised the so-called ***guinea-and-feather experiment***. He allowed a guinea and feather to fall in air and showed that the guinea reached the ground first. The cause of this apparent exception to the teachings of Galileo was discovered by performing the experiment in an exhausted tube. Under such conditions the feather and guinea fell together, for there was then no resisting medium to retard the motion of the feather.

The Acceleration of Falling Bodies.—All bodies fall towards the earth with a constant acceleration g,[1] which is equal to 32 ft. sec.$^{-2}$, or 981 cm. sec.$^{-2}$ in these latitudes. The value of g varies at different places because the earth is neither a true sphere, nor is it homogeneous. In addition, the earth's surface is not smooth, so that the variation of the intensity of gravity with altitude always has to be considered.

When a body moves under the influence of gravity its motion is determined by the equations,

$$\begin{aligned}v &= u \pm gt \\ s &= ut \pm \tfrac{1}{2}gt^2 \\ v^2 &= u^2 \pm 2gs.\end{aligned}$$

[1] 'g' is now termed the *intensity of gravity*—cf. p. 27.

The plus sign is used for falling bodies, the negative for those which rise, and the distance s is considered positive when it is measured vertically upwards. [For bodies starting from rest, $u = 0$.]

Momentum and Force.—Bodies only move relatively to their surroundings if they are acted upon by some ***external agency,*** and by experience we know that it is more difficult to move some bodies than others. This is because the bodies have different masses, where ***mass*** is defined as the quantity of matter in a body. To this statement must be coupled the idea that two masses are equal if, when moving with equal velocities but in opposite senses, they are reduced to rest after a collision in which there is no rebound. The external agency which is capable of imparting motion to a body is called ***force.*** Now when a force acts on an object it cannot increase the mass of the body, and yet we know that the larger the force which acts on a body, the more difficult it becomes for the motion to be arrested. It is the ***momentum*** of the body which has been increased by the larger force, and it continues to be increased by the force during such time that the latter is operative. The momentum I is defined as the product of the mass, m, of the body and its velocity v, so that $I = mv$.

Newton's Laws of Motion.—

LAW I.—***Every body continues in its state of rest or uniform motion in a straight line, unless impressed forces are acting upon it.***

LAW II.—***Change of momentum per unit time is proportional to the impressed force, and takes place in the direction of the straight line along which the force acts.***

LAW III.—***Action and reaction are always equal and opposite.***

GALILEO discovered the first two laws quoted above towards the end of the sixteenth century, whilst the third was known to HOOKE, HUYGHENS and WREN. The three laws were formally stated by Newton in his ***Principia*** in 1686.

A formal proof, analytical or experimental of these laws is not possible, but on them is based the whole system of dynamics, including astronomy. Since the results obtained and the predictions made by astronomers are in good accord with facts, it becomes difficult to imagine that the laws on which their arguments finally depend are erroneous.

Force.—Newton's second law provides us with a means of measuring forces. The proportionality implied in the law may be made into equality by an appropriate choice of units. If the

velocity of a body changes from v_1 to v_2 in time t, the force F is given by

$$F = \frac{m(v_2 - v_1)}{t} = ma \qquad \left[\because a = \frac{v_2 - v_1}{t}\right],$$

only when this particular choice of units has been made; the force then is equal to the change of momentum per unit time.

Units of Force.—When the mass of a body is given in pounds, and the acceleration in feet per second per second, the force is expressed in ***poundals.*** The absolute unit of force in the f.p.s. system is the ***poundal,*** which is defined as ***that force which, acting on a body of mass 1 lb., will impart to it an acceleration of 1 ft. sec.$^{-2}$.*** In the c.g.s. system the absolute unit of force is the dyne, which is ***that force which, acting on a mass of 1 gm., will impart to it an acceleration of 1 cm. sec.$^{-2}$.***

Mass and Weight.—Whenever a body of mass m moves under the influence of gravity it acquires an acceleration g. The magnitude of the force which causes this motion is mg, and it must be attributed to the attraction which the earth has for all matter. This force is called the ***weight*** of the body. We can therefore find the ***weight*** w, of a ***mass*** of 1 lb. It is $w = 1 \times g = 1 \times 32 = 32$ poundals. Hence a mass of 1 lb. has a weight of 32 poundals, by which statement it is to be understood that the force due to gravity on a 1-lb. mass is 32 poundals.

Engineers find that the poundal is too small a unit of force for practical purposes and so choose one which is 32 times as large. This is the pound-weight unit—abbreviated to 1 lb.-wt. The statement that a force of 6 lb. acts on a body has no meaning. The idea which it is intended to convey is that the force is equal to that which is operative on a 6-lb. mass due to gravity, so that its magnitude is (6×32) poundals or $\left(\frac{6 \times 32}{32}\right) = 6$ lb.-wt., the correct statement being that a force of 6 lb.-wt. acts on the body.

Absolute and Gravitational Units of Force.—A poundal and a dyne are termed ***absolute units of force,*** since their values are independent of g, the acceleration due to gravity, a quantity which varies at different places. A pound-weight (1 lb.-wt.) and a gramme-weight (1 gm.-wt.) are called ***gravitational units of force*** since they depend on the value of g.

The Intensity of Gravity.—The force due to gravity acting on a small mass δm near to the earth's surface is $\delta m . g = \delta F$ (say). The force per unit mass is therefore $\frac{\delta F}{\delta m} = g$. From analogy with

the definitions of the strengths of electric and magnetic fields, g is termed the ***intensity of gravity.***

Newton's Third Law of Motion.—Let us examine the statement 'Action and reaction are always equal and opposite' in more detail.

When a book rests on a horizontal table, the action or thrust of the book on the table is equal to the reaction or thrust of the table on the book. In this instance, the action and reaction must be equal and opposite, for if not, motion would ensue. This example, dealing with bodies at rest, presents no difficulty. But Newton's third law of motion postulates that action and reaction are always equal, i.e. even when the two bodies move. Newton considers the particular instance of a horse drawing a cart, 'If action and reaction are equal and opposite, how is it that the horse and cart move forward?' is a question not infrequently asked.

Before attempting to solve this difficulty, which is often a very real one, it must be emphasized that in attempting to solve any mechanical problem, the first essential thing is to fix upon the system whose rest or motion is to be discussed. The system may comprise one body, several bodies, or many bodies, but the system must be clearly defined before the solution is attempted.

In the present problem three possibilities for discussion are as follows :—(i) the motion of the cart, (ii) the motion of the horse, and (iii) the motion of the horse and cart together. If we begin with the first then we must imagine the cart to be isolated from the horse and all other objects by some imaginary closed surface. The forces acting on the cart are

(*a*) an attraction due to the earth—this is called the weight of the cart;

(*b*) an upward thrust (the resultant of the thrusts at the points where the wheels are in contact with the ground).

But by the third law of motion to each of the forces (*a*) and (*b*) equal and opposite forces are exerted by the cart on the earth. These are no concern at present, however, for we have agreed to discuss the motion of the cart, and accordingly have isolated it and have to consider the forces acting on the cart only.

Now the forces (*a*) and (*b*) must be equal and opposite, for if not, the cart would rise or fall according as the thrust of the earth on the cart were greater or less than the attraction of the earth on it. Since there is no motion of the cart in a vertical direction these forces balance, and we do not have to consider them further.

(*c*) Let T be the tension in the traces.

(*d*) Let R be the force due to friction, air resistance, etc. Then if $T > R$ the cart will move forward, its acceleration being

$$\frac{T - R}{M_1} = a_1 \text{ (say)}$$

where M_1 is the mass of the cart.

Similarly, if the motion of the horse is discussed, we shall find that the external forces (not balanced) acting on him are

(i) the tension in the traces (acting backwards),

(ii) the horizontal component, F, of the thrust of the earth on the horse.

If $F > T$ the horse moves forward with an acceleration a_2 given by

$$a_2 = \frac{F - T}{M_2}$$

where M_2 is the mass of the horse.

Since $a_1 = a_2$, the common acceleration of the horse and the cart is

$$\frac{F - R}{M_1 + M_2}$$

If we consider the horse and cart together, the unbalanced forces are F and R, the acceleration being

$$\frac{F - R}{M_1 + M_2}$$

as before. In this instance it is not necessary to consider the tension, since it is now only an internal reaction between two parts of the system; this cannot affect the motion any more than do the intermolecular forces in the horse and cart themselves.

Example.—A mass of 15 lb. is pulled along a horizontal table by a light inextensible string passing over a smooth pulley and carrying a mass of 1 lb. Find the tension (T) in the string, and the acceleration (a) of the system.

Consider the forces acting on the 15 lb. mass. They are :—

(i) In a vertical direction the weight (15×32 poundals) acting vertically downwards which is balanced by the upward reaction of the table.

(ii) In a horizontal direction there is the tension T (poundals). Hence

$$T = 15a.$$

Now consider the forces on the 1 lb. mass.

(i) There are no forces in a horizontal direction.

(ii) Vertically, the weight (1×32 poundals) acts downwards, and the tension T acts upwards.

The resultant force downwards is

$$32 - T$$

and since this acts on a mass of 1 lb., we have

$$32 - T = 1 \times a.$$

Solving these equations

$$a = 2 \text{ ft.sec.}^{-2} \text{ and } T = 30 \text{ poundals.}$$

Example.—Two masses, m_1 and m_2 ($m_1 > m_2$), are connected by a light string passing over a smooth pulley. Discuss the motion and find the tension, T, in the string.

Let a be the acceleration; consider the mass m_1. Resolving vertically, the downward force is $m_1g - T$, and this acts on a mass m_1.

Similarly, for the mass m_2, the upward force is $T - m_2g$.

Hence

$$a = \frac{m_1g - T}{m_1} \text{ or } \frac{T - m_2g}{m_2}.$$

$$\therefore\ a = \frac{(m_1 - m_2)g}{m_1 + m_2}, \text{ and } T = \frac{2m_1m_2}{m_1 + m_2}g.$$

Example. A cage of mass 0·5 ton is drawn up a mine shaft by a rope passing over a smooth pulley at the top. The pull is constant and the cage moves through a distance of 30 ft. in 6 sec. from rest. If the mass of the men inside the cage is 1·5 tons calculate the tension, T, in the rope and the thrust, F, on the floor of the cage.

$$\text{Acceleration} = \frac{2s}{t^2} = \frac{2 \times 30}{36} = \frac{5}{3} \text{ ft.sec.}^{-2}$$

The upward pull on the cage, if T is expressed in ton.-wt., is

$$(T - 2)g = 2 \,.\, \tfrac{5}{3}.$$

$$\therefore\ T = 2{\cdot}10 \text{ ton.-wt.}$$

Consider now the vertical forces on the men; there is F, the upthrust of the floor on them, and their weight $1{\cdot}5g$ downwards. Hence

$$(F - 1\tfrac{1}{2})g = \tfrac{3}{2} \,.\, \tfrac{5}{3} = \tfrac{5}{2}. \qquad \text{[F is expressed}$$

$$\therefore\ F = 1{\cdot}58 \text{ ton.-wt.} \qquad \text{in ton.-wt.]}$$

Atwood's Machine.—If an attempt be made to determine the acceleration due to gravity by observing the motion of a freely falling body, one will soon realize that the method is not susceptible of any great precision since the time of fall is so short that it cannot be measured directly and accurately. The time of fall may be increased by diminishing the downward forces acting on the body. Atwood's machine, in which this principle is involved, is shown in Fig. 2·6 (*a*). It consists of two columns, AB and CD, supported on a common base fitted with levelling screws so that the instrument may be made vertical. The upper ends of these pillars are joined together by a piece of wood upon which is supported a pulley wheel E carried by four other wheels—in this way the friction is reduced to a minimum. Two equal masses, P and Q, usually in the form of cylinders, are attached to a cord passing over E. The mass Q carries a rider of the shape shown at (*b*). Initially Q and its rider rest on a drop bridge L maintained in a horizontal position. When this bridge is lowered the driving force acting upon the system is mg, where m is the mass of the rider and g is the acceleration due to gravity. The load moving is $(2M + m + \kappa)$ where M is the mass of each cylinder and κ is a term added to represent the

inertia of the pulley system. If the whole moves with acceleration a, then

$$mg = (2M + m + \kappa)a$$

or

$$(2M + m + \kappa) = \frac{mg}{a}.$$

In order to eliminate κ the experiment may be repeated using two other equal cylinders of mass M′ when

$$(2M' + m + \kappa) = \frac{mg}{a'}.$$

Subtracting these two equations we have

$$2(M - M') = mg\left[\frac{1}{a} - \frac{1}{a'}\right].$$

From this equation g can be calculated when a and a' are known. To determine the acceleration of a body it is necessary to measure the velocity at two different times. In the present example the body moves from rest, so that its initial velocity is zero. In order to find its velocity, v, at some other time, let us say when the body has moved to R, a ring is placed at this point so that the rider is automatically removed when it reaches this point, and the system continues to move with constant velocity. This is found by observing the time required for the system to move to the lower platform T. If s is the distance RT and the observed time t, the velocity is $\frac{s}{t}$. If we assume that the acceleration over the distance LR is uniform

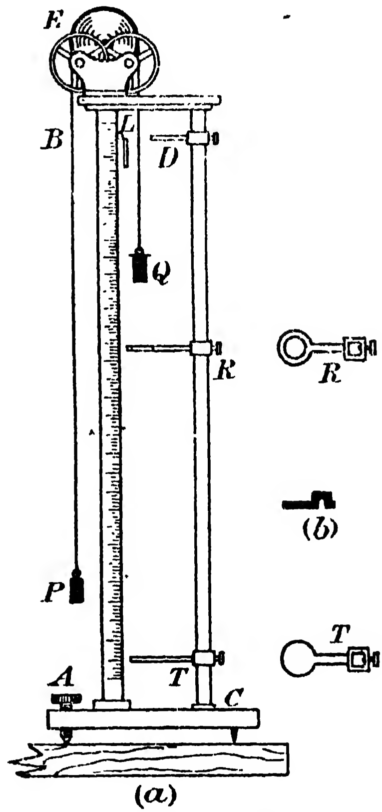


FIG. 2·6.—Atwood's Machine.

$$v^2 = 2ax$$

where $x = $ LR.

$$\therefore a = \frac{v^2}{2x} = \frac{s^2}{2xt^2}.$$

A Modern Form of Atwood's Machine.—CUSSON and JOHNSON have recently designed a form of Atwood's machine in which the method of timing has been considerably improved. The two masses are joined together by a piece of paper in the form of a very light ribbon, Fig. 2·7. This moves past the end of a vibrating arm making a definite number of vibrations per second. The experiment is carried out in exactly the same manner as before, except that the acceleration is now deduced from the wavy trace recorded on the ribbon. An example of such a trace is given in Fig. 2·8. It will be noticed that the initial part of the curve is not very distinct, but by proceeding in the following manner this portion of the curve can be neglected. Imagine a straight line drawn parallel to the edge of the ribbon so that it passes down the centre of the trace. Let A, B, and C be three points at which the wavy line is cut by the straight one and such that the same number of vibrations has been made in the two intervals AB and BC. Let this number be n. If the velocity of the ribbon at A was v_0 and the time taken to make n complete waves t, then

$$s_1 = v_0t + \tfrac{1}{2}at^2$$

where s_1 is equal to AB, and a is the acceleration of the system. Similarly s_2, the distance BC, is given by

$$s_2 = v_1t + \tfrac{1}{2}at^2,$$

where v_1 is the velocity at B. Since this is equal to $v_0 + at$,

$$s_2 = v_0t + at^2 + \tfrac{1}{2}at^2.$$

By subtraction we have

$$s_2 - s_1 = at^2,$$

so that a may be calculated. Care must be taken to see that no portion of the wavy curve used corresponds to the time after the rider has been removed, for the acceleration is then zero. But this portion of the curve may be used to demonstrate that the velocity is then constant.

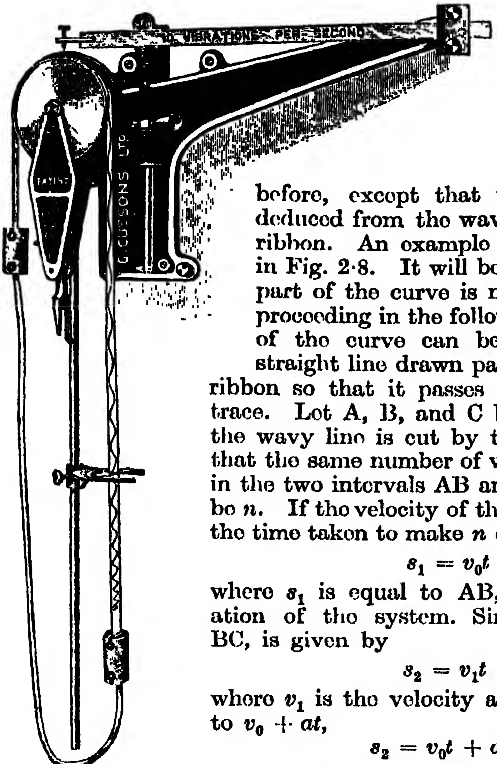


FIG. 2·7.—Modern Form of Atwood's Machine.

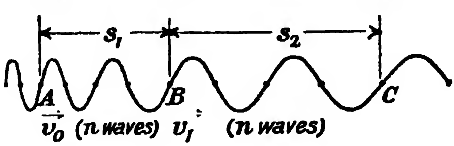


FIG. 2·8.

Welander's Apparatus for Determining 'g.'—A long pendulum, Fig. 2·9, consisting of a thin steel wire and an iron ball B_1 about two inches in diameter, is suspended from the ceiling or wall bracket.

The pendulum is held at an angle to the vertical by an electromagnet, M_1. Another electromagnet, M_2, wired in series with the first, holds a release ball B_2, also of steel. By opening the switch, S, the pendulum and ball B_2 are released simultaneously. On falling to a vertical position the pendulum operates a gate switch, G. A trap switch, T, operates when struck by the falling ball.

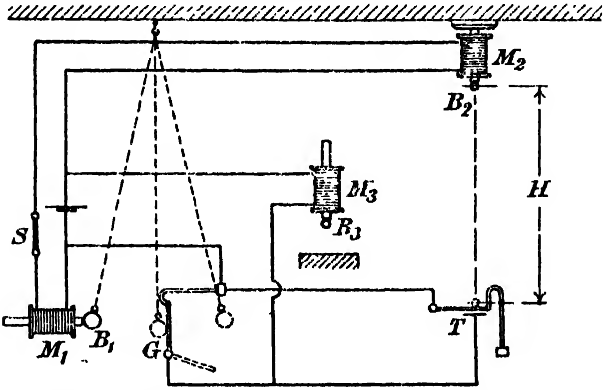


FIG. 2·9.—Welander's Apparatus for Determining 'g.'

To begin the experiment the gate and trap switches are set and the switch S closed. The three electromagnets are now excited so that the pendulum and two spheres may be fixed in position. The trap having been placed at a distance H below B_2, the switch S is opened when the pendulum and ball B_2 are set free together. Let us suppose that the ball B_2 hits the trap switch and closes it before the pendulum strikes the gate G. The electromagnet M_3 will still be excited and B_3 will remain in position. If, however, the gate G is opened before the trap T is closed by impact of the falling sphere, the magnet M_3 is no longer excited and B_3 is released. The trap is therefore moved an inch at a time until two positions of the trap are found such that at one the ball B_3 remains in position while at the other B_3 is set free. By moving the trap 0·1 inch at a time two positions are determined, the lower of which releases the indicator ball while the higher leaves it attached.

Let H be the height through which the ball B_2 falls, and t the periodic time of the simple pendulum used [this must be measured as explained on p. 38]. The release ball strikes the trap after a time $\frac{t}{4}$ so that its velocity is $\frac{gt}{4}$ cm. sec.$^{-1}$ if c.g.s. units are used. But this is equal to $\sqrt{2g\mathrm{H}}$ cm. sec.$^{-1}$. We therefore have

$$\sqrt{2g\mathrm{H}} = \frac{tg}{4}, \text{ or } g = \frac{32\mathrm{H}}{t^2} \text{ cm. sec.}^{-2}.$$

Work.—If a constant force F acts on a body and the point of application of the force moves a distance s along the line of action of the force, work is said to have been done on the body. It is given by $\mathrm{W} = Fs$.

Units of Work.—When a force of 1 poundal is applied so that its point of application moves 1 ft. along the direction in which the force is acting, the work done is ***1 ft.-poundal.*** Engineers use a larger unit called the ***ft. lb.-wt,***[1] which is the work done under conditions similar to the above when the force is 1 lb.-wt.

In the c.g.s. system the absolute unit of work is the ***erg,*** and this is defined as the ***work done when the point of application of a force of 1 dyne moves 1 cm. along its line of action.*** The practical unit is the joule, or 10^7 ergs.

Energy.—If a body is capable of doing work it is said to possess ***energy,*** i.e. the energy of a body is a measure of its capacity for doing work. An agent performing 550 ft. lb.-wt. of work per second does work at a rate of one ***horse-power.*** In the c.g.s. system the power, or rate at which work is done, is often measured in watts, a watt being equivalent to 10^7 erg. sec.$^{-1}$. Electrical engineers find this unit too small for practical purposes so that they generally employ as their unit one which is equal to a thousand watts; it is termed a kilowatt. One horse-power is 0·746 kilowatt. Electrical energy is measured in Board of Trade units, one of which is equal to one kilowatt-hour.

When a mass m is at rest at a height h, it is attracted to the earth by a force mg [its weight]. If it falls to the earth's surface it does work mgh. But $2gh = v^2$, where v is the final velocity of the body, so that the work done is $\frac{1}{2}mv^2$. The body possessed energy when at rest, for it was capable of doing work equal in amount to $\frac{1}{2}mv^2$. The energy which a body possesses in virtue of its position is called its ***potential energy and is measured by the amount of work the body performs in passing from its original position to a standard position, where the potential energy is considered to be zero.*** The potential energy of a body at the earth's surface is taken as zero, whilst the energy associated with its motion is called its ***kinetic energy.***

$$\therefore \text{ Energy} = mgh = \tfrac{1}{2}mv^2.$$

Although the expression $\frac{1}{2}mv^2$ has been obtained for motion under gravity, it is true in general as an expression for the kinetic energy of a mass m moving with velocity v.

Principle of the Conservation of Energy.—From the above it is seen that potential and kinetic energy are mutually convertible —a property which is possessed by all forms of energy, whether they be magnetic, electric, kinetic, etc. When such changes of energy take place ***the principle of the conservation of energy states that the total energy in any system is always constant.***

[1] This is often called a foot-pound, but the notation adopted here leads to less confusion.

Motion in a Horizontal Circle.—If a particle of mass m is moving in a horizontal circle of radius r, with uniform speed v, its velocity cannot be said to be uniform because its direction is continuously changing, and this implies the existence of a force which, in turn, gives rise to an acceleration which can be calculated as follows. Let A and B, Fig. 2·10, be two positions of a particle of mass m rotating in a circle, centre O and radius r. Let v be the speed of the particle, and let us assume that the time required for the body to move from A to B is small. The velocity at A is v and is directed along the tangent AT. At B the velocity is v along BS. If θ is the $\widehat{AOB}$, the velocity at B is equivalent to a velocity $v \sin \theta$ along BD, where BD is parallel to AO, together with a velocity $v \cos \theta$ along BC, where BC is parallel to the tangent AT. If θ is small these are respectively $v\theta$ and v. At A the component velocity along AO was zero so that the change in velocity in this direction is $v\theta$ and this has occurred in a time $\frac{AB}{v}$ or $\frac{r\theta}{v}$. The acceleration is therefore $v\theta \div \frac{r\theta}{v}$, i.e. $\frac{v^2}{r}$. Since the velocity in the direction of the tangent at A does not change by a finite amount, the acceleration along the tangent is zero. The only force acting on the particle is therefore $\frac{mv^2}{r}$ and this is directed towards the centre O; it is called the ***centripetal force***. This force is due to the action of some other body, and since, according to Newton's Third Law of Motion, action and reaction are equal and opposite, it follows that this other body is being pulled by a force which tends to move it from the centre of the circle. This latter is the ***centrifugal force***. Thus, when a stone, attached to one end of a string, is caused to rotate, the pull on the hand of the person performing this experiment is the centrifugal force. The existence of this centrifugal force may also be demonstrated in the following manner. A small container, partly filled with water, is suspended from a string and caused to rotate rapidly in a vertical circle. No water is lost because the velocity is so great that the water exerts a thrust on the bottom of the container which is greater than the weight of the water.

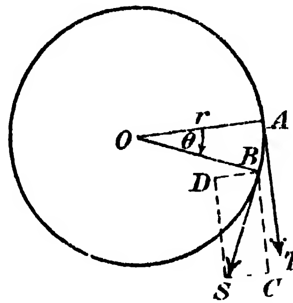


FIG. 2·10.—Motion in a Horizontal Circle.

In chemical laboratories centrifugal force is utilized in the separation of small crystals from the mother-liquors. Dairy farmers also use this force when they separate cream from milk by mechanical means, and in the purification of honey. Dyers are in the habit of rotating their yarns in this way so that they may lose their moisture more readily. In preparing flake nickel for use in Edison storage batteries, the flake is washed and then centrifuged in order to remove the water.

The Banking of Tracks.—Fig. 2·11 represents a truck of mass m moving with speed v on a circular track banked at an angle θ. If G is the centre of gravity of the truck, mg will act vertically downwards along GA whilst the centrifugal force $\frac{mv^2}{r}$ will be operative in a horizontal plane along GB. To prevent the truck from leaving the rails the track must be banked to such an extent that the resultant of these two forces is perpendicular to the track. Under these conditions we have

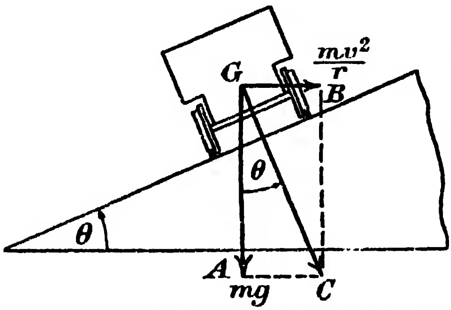


Fig. 2·11.—Truck on a Banked Track.

$$\tan\theta = \frac{\text{AC}}{\text{GA}} = \frac{mv^2}{r} \div mg = \frac{v^2}{rg},$$

where θ is the angle of greatest tilt on the track surface.

If the outer rail is not thus elevated the flanges of the wheels will grind against it in order to create a force sufficient to enable the truck to take the curve at the desired speed.

Simple Harmonic Motion.—Let us imagine that a point is moving with uniform angular velocity ω along the circumference of a circle whose centre is O—cf. Fig. 2·12. If PM is drawn perpendicular to the diameter AOC, M will move to and fro across this diameter as P moves round the circle. The point M is said to execute simple harmonic motion, so that we may say that simple harmonic motion [S.H.M.] is the projection of uniform

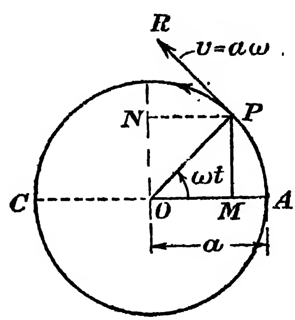


Fig. 2·12.—Simple Harmonic Motion.

motion in a circle upon a diameter of the circle. The distance OA is called the amplitude of the oscillation and the time, T, required for one complete oscillation, i.e. for the point M to move from A to C and back again, is referred to as the period of the oscillation. It is given by the equation $\omega T = 2\pi$, for T is the time required for the moving particle to rotate through an angle 2π.

Let time be reckoned from the instant when the moving point is at A. At time t the moving particle will have reached a point P where $\widehat{AOP} = \omega t$.

Now the velocity of M is equal to the velocity of the point P projected upon AOC, i.e. $a\omega \times \cos \widehat{RPN}$, where $PR = v = a\omega$, and is tangential to the circle at P, and PN is parallel to AO.

$$\therefore v_M = a\omega \sin \omega t. \qquad \because \widehat{RPN} = 90^\circ - \omega t.$$

These equations show the velocity of M is zero at A and C, and reaches a maximum value $a\omega$ at O; the acceleration, on the other hand, is a maximum when M is at A or C, i.e. at either extremity of its path, but it is zero when M is at O [cf. below].

Formulæ for Simple Harmonic Motion.—Let a be the amplitude, while ω is the angular velocity of the point P. The actual speed of P is therefore $a\omega$ so that its acceleration is $a\omega^2$ in the direction PO. The acceleration of M is equal to the component of the acceleration of P parallel to AO, viz. $a\omega^2 \cos \widehat{POM}$. But since $\cos \widehat{POM} = \dfrac{OM}{OP}$, this reduces to $\omega^2 . OM$. We therefore see that the acceleration of M is directly proportional to its distance from O since ω^2 is a constant. When a body moves so that its acceleration is always proportional to its distance from some fixed point in the line of motion and directed towards that point, its motion is said to be simple harmonic.

The periodic time, T, is equal to the time occupied by P moving once round the circle. Now in this time P moves with a speed v a distance $2\pi a$, so that $T = \dfrac{2\pi a}{v} = \dfrac{2\pi}{\omega}$.

Since ω^2 is the acceleration when the particle is at unit distance from the origin, we have

$$T = 2\pi/\sqrt{\text{acceleration for unit displacement}} = 2\pi\sqrt{\frac{\text{displacement}}{\text{acceleration}}}.$$

To Determine the Period of a Simple Pendulum.—A simple pendulum is really a mathematical ideal to which we can only approximate in practice, for it is defined as a heavy particle suspended from a rigid support by a massless inextensible string. The pendulum we have to use consists of a heavy 'bob' suspended by

a light cord. Let A, Fig. 2·13 (*a*), be the 'bob' of mass m, C the point of suspension, so that AC is the string of length l. Let OC be the rest position of the pendulum. Let θ be the $\widehat{ACO}$. Then when the mass is at A the force acting upon it in the direction of the tangent is $mg \sin \theta$, which we may replace by $mg\theta$ if θ is small. The acceleration of A is therefore $g\theta$, i.e. $g \cdot \frac{OA}{l}$, or $\left(\frac{g}{l}\right) \times$ arc OA. Hence, when θ is small, the acceleration is proportional to the dis-

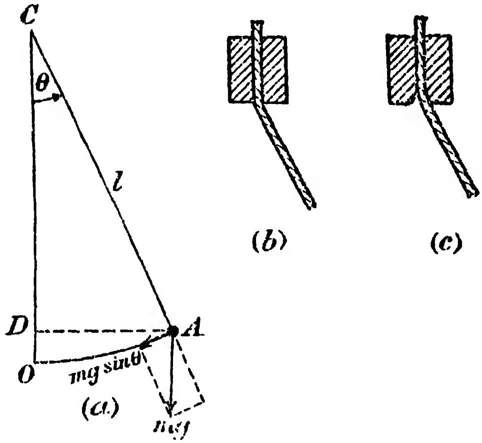


FIG. 2·13.—Simple Pendulum.

tance of A from O so that A will execute a simple harmonic motion, the period of which is

$$T = 2\pi \sqrt{\frac{l}{g}}.$$

To Determine the Acceleration due to Gravity.—The simple pendulum furnishes us with a ready means of finding '*g*.' A lead sphere about one inch in diameter is suspended by a thread held between two blocks as shown at (*b*) so that the length of the pendulum does not vary as it swings. If the lower edges of these blocks are not in the same plane or have become rounded as in (*c*), considerable errors will be introduced. This pendulum is placed in front of some vertical line to indicate the rest position, or it may be viewed with a telescope which is adjusted so that the vertical cross-wire in the telescope coincides with the zero position of the extremity of the pendulum. The pendulum is set in motion and the time of twenty complete oscillations found. The observations are repeated and the mean time calculated. To determine this more accurately the time is noted when the pendulum swings

past its zero position. The swings are ***not*** counted, but after the lapse of some minutes, or previously if the motion is damped considerably, the time is again observed when the pendulum moves through its zero position. The time between these two 'coincidences' divided by the approximate value of the periodic time would be an integer if the observations were not subject to error. The integer nearest to the quotient actually obtained is therefore the number of swings made between the two coincidences; since this is known the period can be calculated. This method of measuring a period should always be resorted to whenever possible.

The length of the pendulum is then changed and the period determined. In this way a series of corresponding values connecting T and l are obtained. Since l is equal to $\frac{gT^2}{4\pi^2}$ it follows that if T^2 [abscissa] is plotted against l [ordinate] the slope of the line will be $\frac{g}{4\pi^2}$; if this is measured g may be deduced [cf. p. 14].

In evaluating the 'slope' of a straight line it must be understood quite clearly that it is not the tangent of the actual angle which the straight line makes with the axis of x, for this depends on the scales used to plot the variables. If A and B are two points on a straight line, and C the point of intersection of straight lines through A and B parallel to the axes Ox and Oy respectively, then the slope of the line is equal to

$$\frac{\text{the quantity represented by CB}}{\text{the quantity represented by AC}}.$$

Motion of a Liquid in a U-tube.—Let l be the total length of liquid [say mercury] contained in a **U**-tube of uniform cross-section. Let m be the mass of liquid per unit length of the tube. If the mercury is depressed in one limb so that it rises an equal distance in the other, on releasing the mercury it will continue to oscillate with a definite period which may be calculated as follows. When the mercury in one limb is at a height x above its zero position the force operative on all the mercury in the tube is equal to the weight of a mercury column of length $2x$, i.e. $W = 2mxg$. Since the total mass of mercury is ml the acceleration, a, at this particular instant is given by

$$2mxg = mla, \text{ or } a = \frac{2gx}{l}.$$

We therefore see that the acceleration is proportional to the displacement x; i.e. the motion is simple harmonic and the periodic time is

$$T = 2\pi\sqrt{\frac{l}{2g}},$$

since $\frac{2g}{l}$ is the acceleration for unit displacement, i.e. the periodic time is the same as that of a simple pendulum of length $\frac{l}{2}$.

Motion of a Body Suspended by a Spring.—We now consider the motion of a heavy body, suspended from a fixed support by a helical spring of negligible mass. Let M be the mass of the body. When the spring is at rest its lower end will be at some definite position. When a small additional mass m is added, let x be the distance through which the lower end of the spring descends. Experiment shows that if a mass m had been removed from the total load carried by the spring the equilibrium position of the lower end of the spring would have been raised by an amount x. It also shows that x is directly proportional to m.

If therefore the load is M and the spring is stretched further by an amount y, the force tending to restore the load to its equilibrium position is the resultant of the weight Mg acting vertically downwards and the upward pull $\left(\mathrm{M} + \frac{m}{x}y\right)g$, viz. $\frac{mgy}{x}$. The acceleration, a, of the mass M will therefore be given by $a = \frac{mg}{\mathrm{M}x}y$, i.e. it is proportional to the displacement y. The motion is therefore simple harmonic with a period T given by

$$\mathrm{T} = \frac{2\pi}{\sqrt{(\text{acceleration for unit displacement})}}$$
$$= 2\pi\sqrt{\frac{\mathrm{M}x}{mg}}.$$

THE THEORY OF DIMENSIONS.

The Dimensions of a Physical Quantity.—It has already been seen that the magnitude of a physical quantity may be expressed in terms of an appropriate unit, i.e. a given quantity is said to be so many times a certain unit. The statement that the length of a particular wall is a metres implies that its length is a times a certain unit of length—the metre. The above statement really consists of two parts—

(i) the pure number or numeric a which is the measure of the quantity in terms of the unit employed,

(ii) the name of the unit.

Now the measure of a physical quantity varies according to the size of the unit employed, but the product of the measure of a physical quantity and the unit employed remains constant. Thus,

2 metre. = 200 centimetre.

If n and n_1 are the measures of a particular physical quantity when the units are $[U]$ and $[U_1]$ respectively, then

$$n[U] = n_1[U_1],$$

i.e.

$$n \propto \frac{1}{[U]},$$

or the measure of a physical quantity is inversely proportional to the size of the unit in which that physical quantity is expressed.

In selecting the units of length, mass, and time the choice is arbitrary. When we have to deal with velocity for example, however, we could still choose a certain velocity as the unit velocity. The unit chosen must satisfy the following requirements :—It must be reproducible and capable of being easily applied. Such a unit of velocity is not easy to find, but the difficulty is overcome in the following way. Suppose that the engine known as the 'Royal Scot,' when travelling at its maximum speed, takes a seconds to pass from one end of the platform of a certain station to the other, the distance between these ends being b cm. The velocity of the engine may then be said to be unity. According to this scheme, the velocity of an object moving 1 cm. in a sec. would be $\frac{1}{b} \times$ (the above unit of velocity). If the distance moved were 1 cm. in 1 sec., the velocity would be $\frac{a}{b}$ units. If an object travelled l cm. in t secs. its velocity in terms of the unit selected would be $\frac{a}{b} \cdot \frac{l}{t}$. In all such expressions the factor $\frac{a}{b}$ occurs. Why not get rid of it by choosing a more suitable unit ? Let the unit of velocity be such that an object moving with unit velocity travels 1 cm. in 1 sec. Then the velocity of a body describing l cm. in t secs. is $\frac{l}{t}$ cm.sec.$^{-1}$. The unit velocity is therefore expressed in terms of the units of length and of time. Such a unit is known as a ***derived unit.***

The unit of velocity thus selected is directly proportional to the unit of length and inversely proportional to the unit of time, for if a body moves 1 metre in a second its velocity will be 100 times that of a body moving with unit velocity, while if it moves 1 cm. in 1 minute, its velocity will be 60 times less. We say that the dimensions of the unit of velocity are 1 in length and -1 in time, a fact represented symbolically as $[L][T^{-1}] = [V]$.

Dimensional Equations.—The interrelationship between the units of length, mass, and time—the so-called ***absolute units***—and a derived unit may be expressed by means of a dimensional equation, where by the statement that the dimensions of a certain

physical quantity are α, β, and γ in length, mass, and time respectively, we mean that the unit in terms of which the quantity is expressed varies as

$$[L]^{\alpha}\,[M]^{\beta}\,[T]^{\gamma},$$

where $[L]$ denotes the unit of length, $[M]$ that of mass, and $[T]$ that of time.

An expression such as $n[L]$ implies that the length of a certain object is n times the unit of length. n is itself a mere number. To discover the manner in which the unit of area depends on that of length let us consider the area of a rectangle whose adjacent sides are b and c. Then its area, a, is

$$\begin{aligned} a[A] &= b[L] \times c[L] \\ &= bc[L]^2. \end{aligned}$$

This equation shows that if, for example, the unit of length is doubled, that of area is quadrupled, i.e. the number expressing the area in terms of the new unit will be reduced to one-fourth of the number expressing the area in terms of the old unit.

Let us consider the dimensions of the unit of density in respect to the dimensions of the three fundamental units. We have

$$\text{density} = \text{mass/volume}.$$

Therefore

$$[D] = [M]/[L]^3 = [M]\,[L]^{-3}.$$

The dimensions of density are therefore one in respect to the unit of mass and -3 in respect to that of length.

Such equations as these are useful in two ways :—

(i) we can express a density given in one system of units in terms of any other possible system of units,

(ii) in any equation in which a number of terms are added together the different terms must be homogeneous as far as their dimensions are concerned, i.e. each term must be expressed in the same units of dimensions. This fact was first pointed out by Fourier, a celebrated French mathematician.

The Dimensions of Some Physical Quantities in Mechanics.—In determining the dimensions of a unit in which a physical quantity may be expressed, it is only necessary to write down any equation, so long as it is valid, connecting this quantity with others whose dimensions are known. The particular equation may be either one applicable to a particular instance or one that is true in general.

(i) *Velocity*. We have

$$s = vt$$

where s is a numeric representing the distance traversed in t seconds

by a body moving with constant velocity v, [t and v are numerics]. Then

$$s[L] = v[V] \, . \, t[T]$$

where $[V]$ denotes the dimensions of the unit in which a velocity is expressed. From the above it follows at once that

$$[V] = [L]\,[T]^{-1}.$$

(ii) *Acceleration.* We have $s = \frac{1}{2}at^2$, so that

$$s[L] = \tfrac{1}{2}a[A] \, . \, t^2[T]^2$$

where $[A]$ denotes the dimensions of the unit in which an acceleration may be expressed. Hence

$$[A] = [L]\,[T]^{-2}$$

(iii) *Momentum* (I). We have

$$I = mv,$$

so that

$$I[I] = m[M] \, . \, v[LT^{-1}]$$

$$\therefore \; [I] = [M][L][T]^{-1}.$$

(iv) *Force* (F). We have $Ft = I$.

$$\therefore \; [F]\,[T] = [MLT^{-1}]$$

$$\therefore \; [F] = [M][L][T]^{-2}.$$

(v) *Work* (W). $W = F \, . \, s.$

$$\therefore \; [W] = [MLT^{-2}][L] = [M][L]^2[T]^{-2}.$$

(vi) *Power* (P). $Pt = W.$

$$\therefore \; [P] = [M]\,[L]^2[T]^{-3}.$$

Some Applications of the Method of Dimensions.—

(i) *The Period of a Simple Pendulum.* This may depend on

(*a*) the mass, m, of the bob,

(*b*) the length, l, of the string,

(*c*) the acceleration, g, due to gravity.

Let us suppose that $t = \kappa m^\alpha l^\beta g^\gamma$, where α, β and γ are the 'exponents of the dimensions,' and κ is the constant.

$$\therefore \; [T] = [M]^\alpha \, . \, [L]^\beta \, . \, [LT^{-2}]^\gamma.$$

Equating like exponents, we have

$$\alpha = 0, \; \beta + \gamma = 0, \; 1 = -2\gamma.$$

$$\therefore \; \gamma = -\tfrac{1}{2}, \; \beta = \tfrac{1}{2}.$$

$$\therefore \; t = \kappa \, . \, \sqrt{\frac{l}{g}}.$$

The constant κ may be determined experimentally.

(ii) *The Natural Frequency* (f) *of a Uniform Stretched Wire.* This may depend on

(*a*) the mass, m, of the wire,
(*b*) its length, l,
(*c*) the stretching force, F.

If $f = \kappa m^{\alpha} l^{\beta} F^{\gamma}$, where α, β, and γ are the 'exponents of the dimensions,' and κ is a constant.

$$\therefore\ [T^{-1}] = [M]^{\alpha}[L]^{\beta}[MLT^{-2}]^{\gamma}.$$

$$\therefore\ 2\gamma = 1,\ \alpha + \gamma = 0,\ \beta + \gamma = 0.$$

$$\therefore\ \alpha = -\tfrac{1}{2},\ \beta = -\tfrac{1}{2},\ \gamma = \tfrac{1}{2}.$$

$$\therefore\ f = \kappa\sqrt{\frac{\text{F}}{ml}}.$$

If μ = mass per unit length of the wire, $m = \mu l$, and

$$f = \frac{\kappa}{l}\sqrt{\frac{\text{F}}{\mu}}.$$

EXAMPLES II

1.—A body, initially travelling with a velocity of 10 ft. sec.$^{-1}$, is observed to be moving with a velocity of 13·1, 15·2, and 16·3 ft. sec.$^{-1}$ at the end of the 1st, 2nd, and 3rd seconds of its motion respectively. Determine the distance traversed.

2.—A train is travelling in a curve of 1 mile radius at a rate of 20 m.p.h.; through what angle has it travelled in 15 secs.?

3.—What do you understand by the term acceleration? A particle has an initial velocity of 14 ft. sec.$^{-1}$. After traversing 100 ft. its velocity is 19·5 ft. sec.$^{-1}$. What is the acceleration, and how long has it been moving?

4.—A body starting from rest is observed to traverse 60 cm. in the 8th second of its motion. What is the acceleration?

5.—A body is thrown upwards with a velocity of 153·2 ft. sec.$^{-1}$. What time elapses before it reaches the highest point in its motion? How high does it rise?

6.—Two buckets, each of mass 7·5 lb., are supported by a thin rope over a smooth pulley and are at rest. A mass of 1 lb. is dropped from a height of 4 ft. into one of the buckets. Calculate the time which elapses before the system has moved through a vertical distance of 10 ft.

7.—Explain what information may be obtained from a graph in which velocity is plotted against time. A train starting from rest accelerates uniformly until it has traversed 1¼ miles; its speed then remains constant for the next 2½ miles when an application of the brakes produces a uniform retardation bringing it to rest after a further ¾ mile. If the whole journey occupies 7½ minutes find the maximum speed in miles per hour. (L.S.C.)

8.—A ball is thrown at an angle of 45° to the horizontal so that at the top of its flight it enters a window 36 ft. above the thrower. Find the speed at which it was thrown and the distance of the wall containing the window from the thrower. (L.S.C.)

9.—Distinguish between *momentum* and *kinetic energy*. Which is conserved during a collision and what happens to the other? A

bullet weighing 30 gm. and travelling at 500 metre. sec.$^{-1}$ embeds itself in a suspended lump of wood of mass 7·47 kilograms. How far will this block have risen above its original position when it reaches the end of its swing? If the length of the suspension is 50 cm. how far will the block have swung in a horizontal direction? (Take: g 1,000 cm.sec.$^{-2}$.] (L.S.C.)

10.—A boy weighing 8 stone and riding a bicycle weighing 21lb. rides up a hill with a gradient of 1 in 21 at 9 ml.hr.$^{-1}$. Assuming that friction is equivalent to a force of 2 lb.-wt. resisting his motion up the hill, find how much work he is doing per second.

11.—Two bodies initially at rest and of mass 10 gm. and 50 gm. respectively are each acted on by a force equal to the weight of a body of mass 4 gm. Compare the times for which these forces must be operative to produce (*a*) the same kinetic energy, (*b*) the same momentum.

12.—Describe the variations of velocity and acceleration of a body moving with simple harmonic motion. If, in a simple harmonic motion, the amplitude of the displacement is 10 cm. and the period 3 seconds, what are the maximum values of the velocity and the acceleration? (B.S.S.C.)

13.—Define two units of force which are in common use. Calculate the force necessary to bring to rest a motor-car weighing 2 tons travelling at a speed of 30 ml.hr.$^{-1}$, in a distance of 20 yds.

14.—Derive an expression for the period of a body moving with simple harmonic motion, in terms of its acceleration and displacement. A vertical U-tube of uniform cross-section contains mercury to a height of 20 cm. If the liquid on one side is depressed, and then released, the mercury oscillates up and down the two sides of the tube. Show that the motion is simple harmonic, and calculate its period.

15.—Define potential energy and kinetic energy, and state the units in which each is measured. A block of wood weighing 500 gm. is allowed to fall down an inclined plane which makes an angle of 30° with the horizontal. After sliding a distance of 20 cm. from rest it is moving with a velocity of 50 cm.sec.$^{-1}$. How much energy has the block lost at this point? What has become of the energy?

16.—Distinguish between *momentum* and *kinetic energy*.

A simple pendulum l metres long has a bob of mass m gm. Derive expressions for the momentum and the kinetic energy of the bob at its lowest point, if the pendulum swings 30° from the vertical.

CHAPTER III

THE ELEMENTS OF STATICS

In the previous chapter it has been shown that whenever a force acts on an object which is not fixed, then that body moves. If the body is to remain at rest it must be acted upon by an equal and opposite force or its equivalent. Under such conditions the body is said to be in ***equilibrium,*** and statics is that branch of physics which studies the properties of bodies in equilibrium. The bodies are supposed to be rigid, homogeneous and not too large, for otherwise the lines of action of all the gravitational forces acting on the individual parts of the body would not be parallel to one another and the problem of determining the line of action of the resultant of such a system of forces is, in general, not capable of solution.

Just as a velocity can be represented by a straight line, so can a force be similarly represented, for this latter has magnitude, direction, and sense.

Resultant of Two Non-Collinear Forces.—If OA and OB, Fig. 2·2, p. 20, represent two forces, F_1 and F_2, the resultant, F, is represented by the diagonal OC, since forces are vectors. Its magnitude is given by

$$F^2 = F_1^2 + F_2^2 + 2F_1F_2 \cos \theta,$$

where θ is the angle AOB.

The Experimental Verification of the Law of Forces.—The experimental arrangement is shown in Fig. 3·1. Three spring balances L, M, and N are supported on hooks, and joined together by means of three pieces of string knotted together at O. The readings of the two balances M and N are observed, and these are a measure of the tensions in the strings. Immediately below the strings a piece of paper is attached to the board which supports the apparatus, and upon this paper straight lines are drawn parallel to the strings leading to M and N. Along these lines distances OA and OB respectively are marked off, their lengths being proportional to the readings of M and N respectively. The parallelogram OACB is then completed, and the tension in L should be proportional to the length of OC whilst the directions OL and OC should be parallel.

Now the reading of the spring balance L measures that force which prevents O from moving when acted upon by the forces in M and N, i.e. the force in L is the ***equilibrant*** of these two other

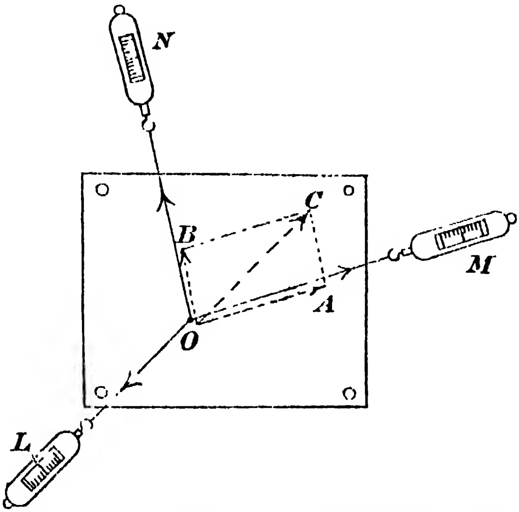


FIG. 3·1.—Verification of the Parallelogram Law of Forces.

forces, whereas the resultant of the forces represented by OA and OB is represented by OC and is equal and opposite to the equilibrant.

Parallel Forces.—When two or more non-collinear parallel forces act upon a rigid body the line of action of the resultant may be found very easily in the following way. If OA and O′A′ Fig. 3·2, represent completely two parallel forces acting on a rigid body, join OO′ and at O and O′ insert two equal and opposite forces OB and O′B′. These will not affect the equilibrium of the body. The two forces at O and O′ are combined according to the parallelogram law, and so we have their resultants OC and O′C′. The lines of action of these two

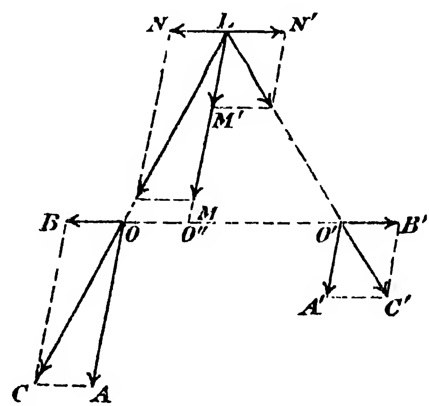


FIG. 3·2.—Graphical Determination of the Resultant of two Parallel Forces.

resultants are produced backwards to meet at L and are there resolved into their components LM, LN, etc. This step is justified, for a force can be represented by a line of suitable length drawn from any point in its line of action. The four forces at L now give a resultant LM + LM′ parallel to the lines of action of OA and O′A′, for the forces LN and LN′ nullify each other. By producing LM to cut OO′ in O″, the point in OO′ through which the line of application of the resultant passes is determined.

Moment of a Force.—Let AB, Fig. 3·3, represent a force F completely, and let OL be the perpendicular from any point O upon AB. Then F . OL is called the moment or torque of the force about O. This moment is represented graphically by twice the area of the Δ OAB, for F . OL is AB . OL which is twice the area of the triangle.

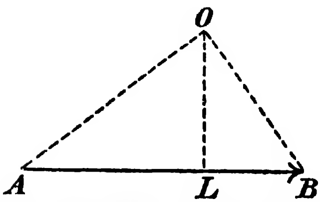


FIG. 3·3.—Moment of a Force.

Couples.—Two equal unlike parallel forces which are not collinear constitute a couple, the moment of which is equal to the magnitude of one of the forces multiplied by the perpendicular distance between the lines of action of the forces.

Centre of Gravity.—If a body is sufficiently *small*, its weight may be regarded as the resultant of the *parallel* forces acting on its constituent particles. It is found that for all such bodies there is some point, not necessarily *in* the body, but which has a definite position with regard to any point in the body taken as reference, and through which the line of action of the resultant of all these parallel forces passes irrespective of the actual position of the body. This point is called the *centre of gravity* [c.g.] of the body. The centre of gravity of any plane body, i.e. a lamina, such as a triangular sheet of metal of uniform thickness, may be found by suspending the body from any point and placing a plumb line immediately in front of the triangle and in such a position that it passes in front of the point of support. Under these conditions the plumb line indicates the line of action of the weight of the suspended body. This direction can then be marked on the triangle. The above procedure is repeated, the lamina being supported from another point. The centre of gravity is then that point at a distance one-half the thickness of the material behind the point of intersection of the lines indicating two positions of the plumb line.

The centre of gravity of a body such as a chair, or bird-cage, is more difficult to find since it is not easy to mark the position

of the plumb line which must invariably be used. It may be done, however, by attaching small pieces of plasticine to the cage and fixing straws therein so that the extremities of the straws touch the plumb line. The extremities of the straws are then joined by a silk thread attached by means of glue. A second determination gives the position of the centre of gravity, for it will be the point at which the two silk threads intersect.

Stable and Unstable Equilibrium.—Whenever a body is in statical equilibrium the resultant force upon it must be zero, but the nature of the equilibrium is not always the same. To illustrate these remarks let us consider the equilibrium of a sphere resting in turn on a concave, a convex, and a flat surface as shown in Fig. 3·4. When the sphere is given a slight displacement from its zero position on a concave surface it tends to return to this position as soon as the constraining force is removed. The equilibrium is said to be ***stable.*** The equilibrium, however, when the sphere rests on a convex surface is ***unstable,*** because if the sphere experiences even a very small displacement it never returns of its own accord to its former position. In the third case when the sphere rests on a flat surface the equilibrium is called ***neutral*** because the body may be at rest at any point on the surface.

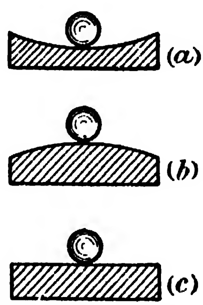


FIG. 3·4.—Types of Equilibrium.

Machines.—As a result of experience man has found that he can work better in some positions than in others; it is always more convenient to pull a rope downwards rather than upwards, and it is generally more convenient to apply a small force through a given distance when it would be impossible, for example, to apply a force ten times as great through one-tenth that distance. Hence it is desirable to have a contrivance for changing (*a*) the point of application, (*b*) the magnitude of a force and (*c*) its direction. ***A mechanical arrangement whereby a force acting at one point is made available at another point under different conditions as regards its magnitude and direction is known as a simple machine.*** If F is the effort or force necessary to be applied to a given machine to overcome the weight W of a load which it carries, $\frac{W}{F}$ is called the ***mechanical advantage*** of the machine.

Another quantity often evaluated in connexion with a machine is its ***velocity ratio.*** This is defined as the ratio of the distance through which the point of application of the effort moves to the

distance through which the point of application of the resistance or weight moves in the same time, i.e.

$$\text{Velocity ratio} = \frac{\text{distance through which F moves}}{\text{distance through which W moves}}.$$

If a machine is perfectly smooth and no energy is used in moving the component parts, the work done against W is equal to the work done by F,

i.e. $$\text{W} \times \text{distance through which W moves}$$
$$= \text{F} \times \text{distance through which F moves.}$$

Hence, $$\frac{\text{W}}{\text{F}} = \frac{\text{distance through which F moves}}{\text{distance through which W moves}},$$

or the mechanical advantage of a machine in which there is no friction and in which the parts have no weight is equal to its velocity ratio. In any actual machine the mechanical advantage must be determined experimentally but in simple machines it is possible to calculate the velocity ratio.

The efficiency of a machine is defined by the equation

$$\begin{aligned}\text{Efficiency} &= \frac{\text{work done by the machine}}{\text{work done by the effort}}\\ &= \frac{\text{load (W)} \times \text{distance load moves}}{\text{effort (F)} \times \text{distance effort moves}}\\ &= \frac{\text{W}}{\text{F}} \div \text{velocity ratio, i.e.}\end{aligned}$$

$$\text{Efficiency} = \frac{\text{mechanical advantage}}{\text{velocity ratio}}.$$

The Principle of Virtual Work.—Mechanical problems, especially those dealing with simple machines, i.e. machines without friction and in which no work is required to move the components, may be solved by a principle first pointed out by STEVINUS in connexion with pulleys. He noticed that when a load of weight mg or W is raised by a cord passing over a single fixed pulley, that the effort is equal to the weight and that the point of application of the effort descends through a vertical distance equal to that through which the weight is raised. In the instance of a single movable pulley, the effort is only one-half of the weight of the load raised, but its point of application moves through twice the distance. Stevinus argued that this principle applied to all simple machines and wrote 'Ut spatium agentia ad statiam patentis, sic potentia patentis ad potentiam agentis,' a free translation of which is 'What is gained in power is lost in speed.' A better statement of this principle is that mechanical advantage is always gained at a proportionate diminution in speed.

In 1717 BERNOULLI, an eminent mathematician, extended the above principle to all cases of equilibrium. He maintained that if any number of forces acting on a body undergo infinitely small displacements consistent with the configuration of the system, then the total work done is zero, i.e.

$$\Sigma F \cos \alpha . \delta s = 0$$

where δs is the displacement of the point of application of F, and α is the angle between F and δs. The necessity for the displacements to be infinitely small follows at once from the fact that if they are finite the system may assume another configuration in which equilibrium is only maintained under conditions different from those for the given system. This principle, the so-called principle of virtual work, will be used in discussing some of the problems which follow.

Since no machine is without friction, etc., the principle of work, as here used, only allows us to calculate $\frac{W}{F}$ on the assumption that the machine is ideal, i.e. it gives us the velocity ratio in all cases but the mechanical advantage only if the machine is ideal.

Levers.—One of the simplest forms of machine is the lever, of which there are three classes according to the position of the

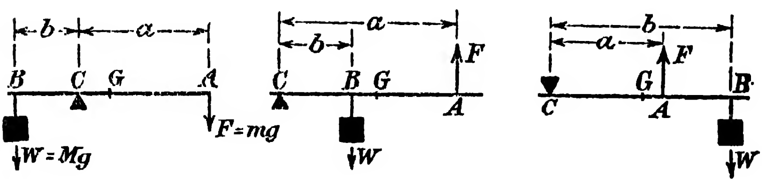


FIG. 3·5.—Levers.

point or *fulcrum* about which they turn. The three classes are shown in Fig. 3·5. In addition to the forces F and W there is the reaction at the fulcrum C, and since there is equilibrium the reaction must be equal to the algebraic sum of F and W. In all three instances when the levers are in equilibrium

$$F . AC = W . BC.,$$

or

$$\frac{W}{F} = \frac{a}{b}.$$

This is the mechanical advantage in so far as friction and the weight of the lever are negligible. The velocity ratio is calculated as follows :—

If the beam rotates through an angle θ, the point of appli-

cation of the effort moves a distance $a \sin \theta$ while that of the load moves a distance $b \sin \theta$, i.e. the velocity ratio is a

In the above it has been assumed that friction and the weight of the lever are negligible. If the weight is W_1, and acts at G, the condition for equilibrium in the first class is

$$W \cdot BC = W_1 \cdot GC + F \cdot AC.$$

Similar expressions are easily written down for the other two classes of lever.

The Balance.—The physical or analytical balance, the main features of which are shown in Fig. 3·6, consists of a light but rigid metal beam, supported so that it may rotate in a vertical plane

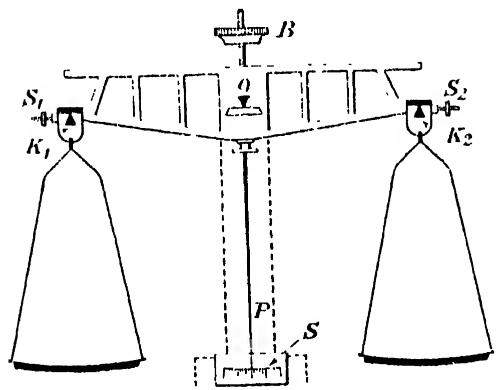


FIG. 3·6.—Main Features of a Physical Balance.

about a horizontal axis vertically above its centre of gravity. Pans are suspended from the extremities of the above beam and turn freely about axes parallel to its axis of rotation. The beam has three agate knife-edges: the central one, O, edge pointing downwards, supports the beam when it rests upon an agate plate attached to the pillar (indicated by the broken lines) of the balance, while the outer ones, edges pointing upwards, carry the agate plates to which the scale-pan supports are fixed. The whole is enclosed in a glass case to protect it from air currents and rapid temperature changes, the air within the case being kept dry by means of some desiccating agent such as concentrated sulphuric acid. When the balance is not in use the knife-edges are shifted slightly so that they are not in contact with the agate plates: in this way the knife-edges are kept in working order for longer periods. The position of the balance beam is defined by a long metal pointer, P, rigidly attached to it, its length being normal to the line joining the outer

knife-edges. The observation of the motion of the pointer, and consequently that of the beam, is facilitated by means of a scale S fixed to the pillar of the balance. Small masses, S_1 and S_2, capable of moving on screws attached to the beam, enable the balance to be adjusted. When the balance is in proper adjustment the pointer should swing through equal distances on either side of the central mark of the scale S when the line K_1K_2 is horizontal. The bob B, also moving on a screw attached to the beam, permits the position of the centre of gravity of the beam and its attachments to be raised or lowered with respect to the central knife-edge O.

Simple Theory of the Static Equilibrium of an Equi-arm Balance.—In principle the common balance is simply a lever of

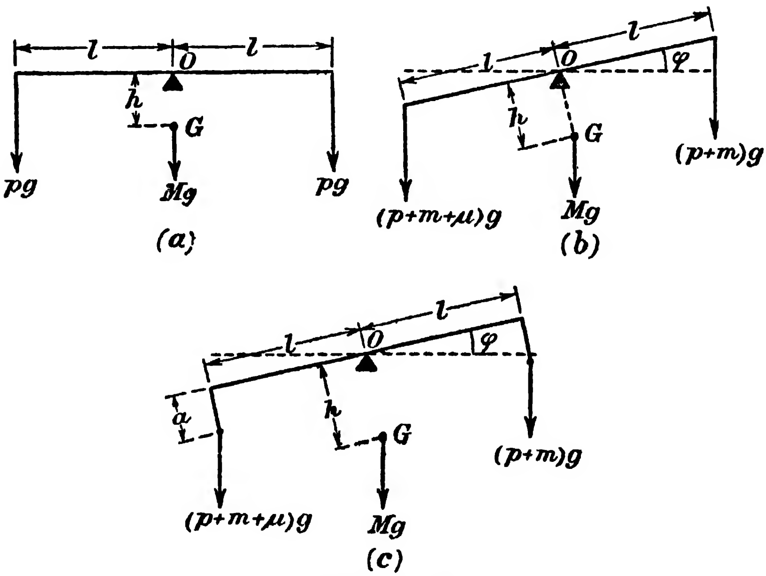


FIG. 3·7.—Sensitivity of a Balance.

the first class in which the two arms are equal and the central knife-edge O, Fig. 3·6, is the fulcrum. As a first approximation let us assume that the effective lengths of the arms are equal and invariable and that the three knife-edges are coplanar, horizontal and parallel to one another, as suggested in Fig. 3·7 (*a*). Let M be the mass of the beam, and its attachments, while G is their common centre of gravity. Further, let p be the mass of each pan and suppose that a mass m is placed in one pan, while $(m + \mu)$ is the mass in the other pan, where μ is a small mass. Let l be the length of each arm and h the distance OG. When the beam is loaded as above, let it take up an equilibrium position inclined

at a small angle ϕ to the horizontal—cf. Fig. 3·7 (*b*). Then, if g is the intensity of gravity, by taking moments of forces about O, we have

$$(p + m + \mu)gl \cos \phi = Mg\, h \sin \phi + (p + m)gl \cos \phi$$

$$\therefore \mu l \cos \phi = Mh \sin \phi$$

$$\therefore \tan \varphi = \frac{\mu l}{Mh}.$$

Since ϕ is small, $\tan \phi$ may be replaced by its circular measure ϕ, so that $\phi = \frac{\mu l}{Mh}$.

Now the ***sensitivity*** of a balance is defined as the change in ϕ caused by increasing μ by a given small amount, usually taken as 10^{-3} gm., i.e. if $\delta\phi$ is the change in ϕ thus caused,

$$\phi + \delta\phi = \frac{(\mu + 10^{-3})l}{Mh},$$

or
$$\delta\phi = \frac{10^{-3}l}{Mh}.$$

Thus, for the ideal balance, in which the knife-edges are coplanar, the sensitivity is independent of the load in the pans. From the above expression for the sensitivity it follows that, for a given balance, it may be varied by altering h, i.e. the position of G with respect to O. This, as already mentioned, is effected by means of the bob B, Fig. 3·6.

When the knife-edges are not coplanar let the outer edges, when the beam is at rest, be at a depth a below the horizontal plane through O—cf. Fig. 3·7 (*c*). Then, as before,

$$(p + m + \mu)g(l \cos \phi - a \sin \phi) =$$
$$Mg\, h \sin \phi + (p + m)g(a \sin \phi + l \cos \phi),$$

i.e. $\cos \phi[(p + m + \mu)l - (p + m)l] =$
$$\sin \phi[Mh + a(p + m) + a(p + m + \mu)].$$

Since $\phi \to 0$, $\cos \phi \to 1$ and $\sin \phi \to \phi$.

$$\therefore \phi \simeq \frac{\mu l}{Mh + a[\mu + 2(p + m)]}$$

$$\simeq \frac{\mu l}{Mh + 2a(p + m)} \quad [\because \frac{\mu}{p} \to 0].$$

The sensitivity is therefore given by

$$\delta\phi = \frac{10^{-3}l}{Mh + 2a(p + m)}.$$

Thus, if the three knife-edges are not coplanar, the sensitivity decreases with increasing load: since all beams are deformed

slightly by the load, so that the knife-edges never remain coplanar when the balance is loaded [in fact a varies with the load, also], the sensitivity of all balances decreases with the load.

If a balance is to be classified as a good one, it must possess the following characteristics :—

(*a*) Its indications must be reliable; i.e. the beam must be horizontal when equal masses are placed in the two pans. This is secured by making the arms exactly equal in length and mass; the suspended pans must also be of equal mass.

(*b*) The balance must be sensitive, i.e. a small difference between the two masses compared must cause an appreciable deviation of the beam from its zero position, i.e. ϕ must be relatively large. This is obtained by making M and h small. Hence the beam must be long and light, and have its centre of gravity near to O.

(*c*) A good balance must be stable, i.e. it must not suffer any change in shape, e.g. by bending of the beam, etc. For this reason the sensitivity cannot be increased indefinitely, for such a condition can only be attained by using a light beam, whereas the beam must be fairly massive if it is to be rigid. Evidently these conditions are at variance and, in practice, a compromise must be effected.

(*d*) The period of swing should be short, so that 'weighings' may be made rapidly—unfortunately this implies a less sensitive balance, so that again a compromise is made.

(*e*) The balance must be stable, i.e. when the balance is in equilibrium it must return to its zero position after being deflected.

(*f*) In addition the knife-edges [which, it must be noted, are always fixed to the beam] should be parallel and lie in the same horizontal plane. This latter condition is essential if the sensitivity of the balance is not to vary with the load in the pans.

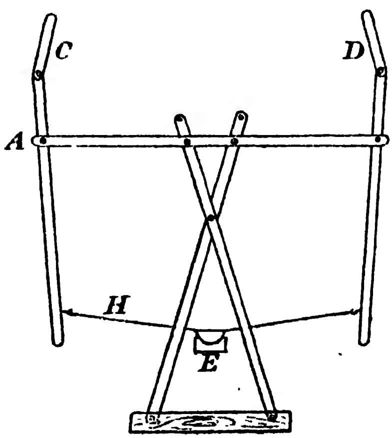


FIG. 3·8.—A Micro-Balance.

A Micro-Balance.—Small masses, such as drops of liquid absorbed in bits of filter paper, or small quantities of powder, may be estimated by means of a micro-balance shown in Fig. 3·8. A light thread, H, has its extremities attached to the lower ends of a pair of nearly vertical rods 2 ft. apart, each of which is pivoted at a point just above its centre of gravity. Thus, a very light load suspended at the middle

of the thread causes a considerable depression of that point. The apparatus is calibrated by observing the depression for a known load. The contrivance can be constructed out of 'Meccano' parts. The friction at the fulcrums A and B is reduced by using short glass tubes as the supports for the uprights. By making the uprights in two parts as shown and moving the upper portions C and D the sensitivity may be altered considerably. The 'pan' of the balance consists of a small circular disc bent across one of its diameters so as to form a clip which can be suspended from the thread, as at E. Small objects can then be supported between the jaws of this clip. Such an instrument as this has many uses, especially in the study of bacteriology. It was originally designed for use in Flanders, during the war of 1914–1918.

The Single Movable Pulley.—In this very simple type of machine a string, fastened at one end to a beam, passes round a pulley, K, Fig. 3·9, carrying a load of weight W. The portions of the string passing round the movable pulley are parallel to one another. The effort, or force, F, necessary to raise the load, is applied at the free end of the string which, for convenience, may

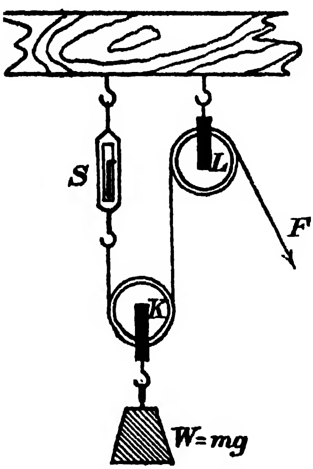


FIG. 3·9.—A Movable Pulley.

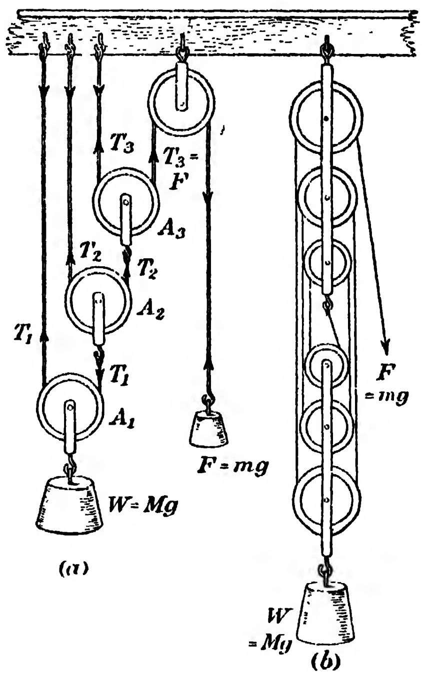


FIG. 3·10.—Systems of Pulleys.

pass round a fixed pulley, L. To determine the force required to maintain W in equilibrium a spring balance, S, is placed as indicated in the diagram. From observations made with such an apparatus it is soon realized that, if friction and the weight of the pulley be neglected, the tension in the string, which is measured by S, is one half the weight of W; this means that one half the

load is supported by the string attached to the beam and the second half by the string passing round the fixed pulley—the free end of this string may be held in any convenient direction. The fact that a movable pulley-wheel with parallel strings reduces by one half the effort required to raise an object is a principle which may always be applied to such pulleys when they are free to move.

Of the various systems of pulleys with parallel strings, and the ways in which pulleys may be combined to form a machine, only two will be considered; they are shown in Fig. 3·10 (*a*) and (*b*).

In the ***Archimedean or First System of Pulleys,*** Fig. 3·10 (*a*), a separate string passes round each pulley. If the load W ascends a distance x, the string round A_1 is shortened by an amount x on each side, so that A_2 moves a distance $2x$. Similarly A_3 moves a distance $2 \times 2x = 2^2x$ and the point at which F is applied descends a distance 2^3x. The velocity ratio is therefore $2^3 = 8$. In the case of n movable pulleys the velocity ratio is 2^n.

To calculate the mechanical advantage of the system in the absence of friction and neglecting the weights of the pulleys we make use of the fact that the work done on W is equal to the work done by F, i.e. if there are three movable pulleys

$$W \,.\, x = F \,.\, 2^3x, \text{ or } \frac{W}{F} = 2^3.$$

Alternatively, let the tensions in the different strings be as shown. Then

$$T_1 = \tfrac{1}{2}W,\ T_2 = \tfrac{1}{2}T_1,\ T_3 = \tfrac{1}{2}T_2 = \tfrac{1}{8}W.$$

Since $T_3 = F$, $\dfrac{W}{F} = 8$, as above.

In the case of n movable pulleys the mechanical advantage, that is $\dfrac{W}{F}$, is 2^n.

In the above argument the weight of the pulleys has been neglected. Suppose that there are three movable pulleys, each of weight w. Then $T_1 = \frac{1}{2}(W + w)$;

$$T_2 = \tfrac{1}{2}[\tfrac{1}{2}(W + w) + w] = \tfrac{1}{4}W + \tfrac{3}{8}w;\ \ T_3 = \tfrac{1}{2}[\tfrac{1}{4}W + \tfrac{3}{8}w + w]$$

$$\therefore F = T_3 = \tfrac{1}{8}W + \tfrac{11}{16}w.$$

$$\therefore \frac{W}{F} = 8 - \frac{11}{2}w\left(\frac{1}{F}\right).$$

Calling $\frac{W}{F}$, the mechanical advantage, y, and $\frac{1}{F} = x$, we have

$$y = 8 - \frac{11}{2}wx.$$

This equation suggests that if corresponding values of x and y are plotted, the resulting graph will be a straight line whose slope is $-\frac{11}{2}w$. Hence w can be determined.

In the ***Second or Common System of Pulleys*** there is only one continuous string and this passes round all the pulleys, its one end being fixed to the upper support, and the pull F applied at the other extremity [cf. Fig. 3·10 (*b*)]. If the load rises a distance x, the amount of rope 'set free' is $6x$, since each string supporting the lower block is shortened by an amount x. To keep the string taut the point of application of the effort must descend a distance $6x$: the velocity ratio is therefore 6. When there are n strings supporting the lower sheave pulley block the velocity ratio is n.

In actual practice the pulleys in each block in the common system are all concentric so that the two blocks can be drawn nearer together. The great disadvantage of these systems is that a long length of rope is required; this is avoided in the differential pulley.

Weston's Differential Pulley.—In this system the rope is replaced by an endless chain, slip being prevented by depressions in the grooves of the pulleys, and into these depressions fit the links of the chain. The system is represented in Fig. 3·11, in which the two pulleys of the upper block move

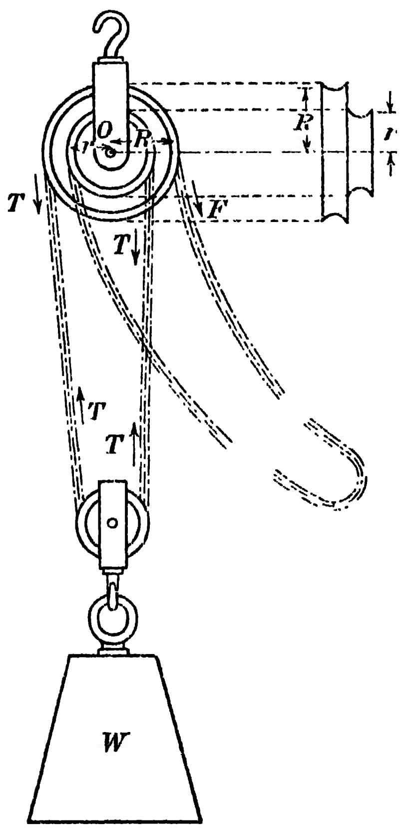


FIG. 3·11.—Weston's Differential Pulley.

as one round a common axis. The effort F is applied as shown. If W is the load and T the tension in the string, the necessary condition for the equilibrium of the load, on the assumption that the machine is an ideal one, is $W = 2T$, whilst, by taking moments of forces round O, the relation

$$F \cdot R + T \cdot r = T \cdot R$$

is obtained. Hence

$$F = T \cdot \frac{R - r}{R} = \frac{1}{2}W \cdot \frac{R - r}{R}.$$

The mechanical advantage, $\frac{W}{F}$, is therefore $\frac{2R}{R - r}$, and since the machine has an efficiency of 100 per cent, this fraction is also the velocity ratio.

The value for this ratio can be found for a differential pulley, even when it is not an ideal system, as follows :—

Suppose that the upper pulley block makes one complete revolution and that W rises. The length of chain wound in $= 2\pi R$, while that let out $= 2\pi r$. Therefore the length of chain actually supporting the lower pulley and W is shortened by an amount $2\pi(R - r)$, i.e. W rises a distance $\pi(R - r)$. Since the point of application of F descends a distance $2\pi R$, the velocity ratio is $\frac{2R}{R - r}$. Hence in this non-ideal machine the efficiency is

$$\frac{W(R - r)}{2FR}.$$

The Inclined Plane.—When a body S, Fig. 3·12, rests on a smooth inclined plane it is acted upon by two forces, the weight of the body acting vertically downwards, and the reaction of the plane on the body which is normal to the surface. The body will therefore move under the influence of the resultant of these forces unless it is constrained by some other force. The two cases which we shall study are when this third force, F, is either parallel to the line of greatest slope in the plane, or to the base of the plane—cf. Fig. 3·12 (*a*) and (*b*). The velocity ratio is $\frac{AB}{BC}$, i.e. cosec θ in the first instance and $\frac{AC}{BC}$, i.e. cot θ in the second. These expressions are valid independently of whether friction is present or not. For any actual plane, as indeed for any actual machine, the mechanical advantage $\frac{W}{F}$ must be determined experimentally. If the plane is smooth the mechanical advantage may be calculated as follows. To determine the magnitude of the effort F required to hold a

body of weight W on a smooth plane whose inclination is θ and when F is parallel to the line of greatest slope AB in the plane we resolve the forces acting on S along AB. This gives

$$mg \cos\left(\frac{\pi}{2} - \theta\right) = \mathrm{F},$$

i.e.

$$mg \sin \theta = \mathrm{F}.$$

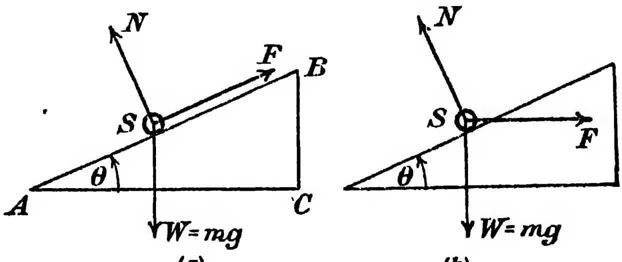


FIG. 3·12.—The Inclined Plane.

Similarly, by resolving forces perpendicular to the plane, we get the normal reaction, N, of the plane on the body, viz.,

$$\mathrm{N} = mg \cos \theta.$$

In these equations it must be remembered that if m is expressed in pounds, F and N are in poundals. The more usual practice is to express the weight mg as W lb.-wt., when the above equations become

$$\mathrm{W} \sin \theta = \mathrm{F} \text{ etc.},$$

where F and N are now measured in lb.-wt. Since $\frac{\mathrm{W}}{\mathrm{F}}$ is the mechanical advantage of the system it follows that this is equal to cosec θ in this instance; in the second it can be shown to be cot θ.

The Screw.—If a triangle PQR, Fig. 3·13, in which $\widehat{\mathrm{QPR}}$ is equal

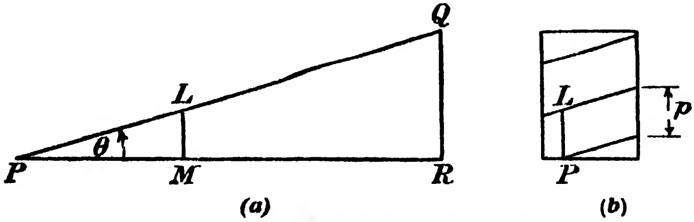


FIG. 3·13.—The Principle of a Screw.

to θ, is constructed out of thin paper or aluminium foil and wrapped round a right circular cylinder so that the base PR remains in a plane perpendicular to the axis of the cylinder the trace of the hypotenuse on the surface of the cylinder is a spiral. Let LM be a line perpendicular to the base of the Δ PQR such that when the paper

is round the cylinder the point M coincides with P, and L is vertically above P. Then, regarding the trace of the edge PQ as the thread of the screw, LM is the pitch of the screw. The $\widehat{QPR}$ is called the angle of the screw and it is clear from the diagram that

$$\tan \theta = \frac{LM}{PM} = \text{pitch of screw} \div \text{circumference of cylinder.}$$

Actual screws differ from this ideal screw in that they always have a protuberant thread of metal or wood, etc. This enables the screw to work in a nut, but of course introduces so much friction that the mechanical advantage of a screw never approaches equality with its velocity ratio which we now proceed to obtain.

The Velocity Ratio and Efficiency of a Screw.—Let us suppose that we have a screw working in a nut and that the screw is supporting a load of weight W, as in Fig. 3·14, while a force F, which we assume to be in a horizontal plane, is applied to the end of the arm AB. Now when the arm AB has made one complete revolution the point of application of F has moved through a distance $2\pi r$ where r is the distance of B from the axis of the screw. Under the same circumstances the load W will have moved through a distance p, where p is the pitch of the screw. The velocity ratio is therefore $\dfrac{2\pi r}{p}$.

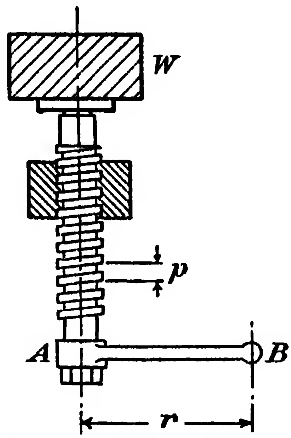


FIG. 3·14.—Mechanical Advantage of a Screw.

If the mechanical advantage, $\dfrac{W}{F}$, is determined experimentally, the efficiency is $\dfrac{W}{F}\cdot\dfrac{p}{2\pi r}$.

Weighing Machines.—The mass of a heavy load may be ascertained with the aid of a weighing machine the principle of which is indicated in Fig. 3·15. It consists of three levers ACD, EK, and LR respectively. The platform upon which the load is placed is attached to the lever EK, whilst the end D of the first lever carries a scale-pan. The fulcrum for the lever EK is not fixed but is attached to the lever LR moving about a fixed fulcrum R. If a load of mass M, and therefore weight Mg, is placed on the platform we may regard its weight as being distributed at the points E and K. The actual distribution will depend upon the position of the load on the platform, but let us suppose that there is a load mg at K so that the load at E is $(M - m)g$. The load mg at K can be replaced by a load

mg/n at L if LR $= n$. KR. Now the load at L may be considered to be acting at A, and may therefore be replaced by a load n times as large at B if AC $= n$. BC, i.e. the equivalent load at B would be mg. But the load at E may be replaced by an equal one at B so that

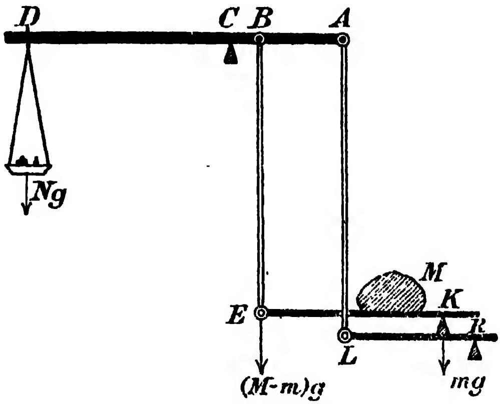


FIG. 3·15.—A Weighing Machine.

the total load at B is now Mg, and this is independent of the actual position of the object on the platform. To measure this load at A the length of the lever CD may be made 10 or 100 times that of AC. When this is done the mass N of the load on the scale-pan is the corresponding fraction of the mass M.

The Common or Roman Steelyard.—This is another machine for determining the mass of a heavy load, and consists of a long non-

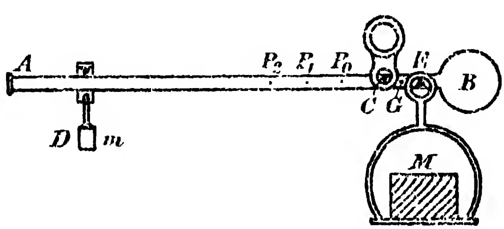


FIG. 3·16.—Common Steelyard.

uniform rod, AB, Fig. 3·16, movable about a fixed fulcrum, C, situated a little to the left of G, the centre of gravity of the bar. A hook, or scale-pan, hung from E, is used to carry the load of mass M, whilst D is a bob of mass m movable along AC. The point at which D must be placed to maintain the steelyard in a horizontal position enables one to determine the mass of the load M.

To calibrate the steelyard let P_0 be the position of D when the load is zero; this point is given by the equation

$$m \,.\, P_0C = \mu \,.\, CG,$$

where μ is the mass of the steelyard. [All masses are expressed in stones, where 1 stone = 14 lb.] When the load in the scale-pan is 1 stone let D be at P_1. The position of P_1 is determined by

$$m \,.\, P_1C = \mu \,.\, CG + (1 \times EC).$$

Subtracting the first equation from this we have

$$m \,.\, P_0P_1 = 1 \times EC.$$

Similarly, when the load is 2 stones the position of D is given by

$$m \,.\, P_0P_2 = 2 \times EC.$$

We see therefore that this instrument may be graduated by engraving marks upon the bar such that their common distance apart is equal to EC/m, the zero division being at P_0 as defined above.

The Danish Steelyard.—This consists of a bar, AB, Fig. 3·17, terminating in a sphere at B. The other extremity of the bar carries a scale-pan to receive the load whose mass is required. The pan is fixed, so that the mass of the load is determined by observing the point in the rod about which it balances. To graduate the steelyard let m be the mass of the whole including the pan, and let G be the centre of gravity. If C is the fulcrum when the load in the pan has a mass M, by taking moments of forces about C we have

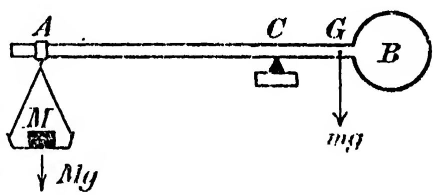


FIG. 3·17.—Danish Steelyard.

$$Mg \,.\, AC = mg \,.\, GC = mg \,.\, (AG - AC).$$

Therefore $$AC = \frac{m}{M + m} \,.\, AG.$$

This equation indicates that to graduate the steelyard it should first be balanced about its centre of gravity, i.e. G is found. Let us further assume that the mass m is 1 stone. The middle point of AG is the fulcrum when the load in the pan is 1 stone. Similarly, when the load is increased to 2 stones the fulcrum must be at a distance $\frac{1}{3}$ AG if the whole is in equilibrium. We therefore see that if the load is n stones the point of balance must be such that

$$AC = \frac{1}{n + 1} AG.$$

Friction.—*Hitherto it has been supposed that the surfaces of* bodies in contact have been perfectly smooth, so that the reaction of one on the other was always directed along the common normal to the surfaces at the point of contact. In practice this condition is only satisfied if there is no tendency for relative motion between the surfaces: when there is such a tendency, forces are called into play and oppose the motion. These forces are due to ***friction*** between the surfaces in contact.

The Laws of Static Friction.—The effects of friction were investigated experimentally by COULOMB in the following manner. A, Fig. 3·18 (*a*), is a board resting on a horizontal table. B is a slider which could be suitably weighted in order to vary the

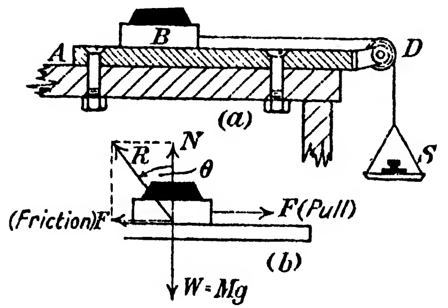


FIG. 3·18.

Coulomb's Apparatus for Investigating the Laws of Static Friction.

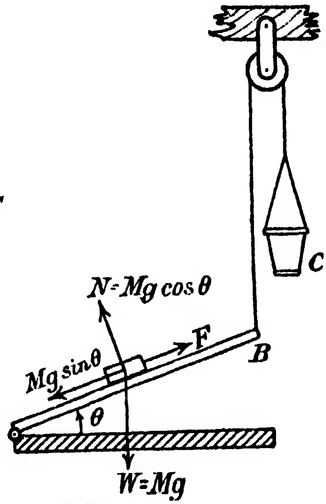


FIG. 3·19.

thrust between the surfaces of A and B in contact. Attached to the slider is a cord passing over a pulley D and carrying a scale-pan, S.

When there is no pull in the cord, the thrust of the board A upwards on the slider balances the weight of the latter and the system is at rest. When S is loaded there is a pull exerted in the string, but, provided that this is lower than a certain limit, no motion ensues. Forces exactly balancing the tension in the cord have been brought into play and resist the motion. These are the ***frictional forces***. It is found that up to a certain limit, in any given instance, just enough friction is called into play to prevent motion. If the pulling ceases, the forces due to friction also cease, for if they did not the body would move. Friction is

a self-adjusting force, for no more friction is called into play than is necessary to prevent motion. The amount of friction which may be exerted between two surfaces in contact is not, however, unlimited, for if the pull in the string is increased gradually a stage is finally reached when the body just begins to move; the friction is said to have reached its ***limiting value***, and if the pull is further increased the slider is accelerated.

From experiments carried out on the lines suggested above, Coulomb established the following facts:—

(i) The limiting friction is independent of the area of contact between the surfaces so long as the thrust between them is unchanged.

(ii) The limiting frictional force, or limiting friction, is directly proportional to the normal thrust between the surfaces in contact, when the materials and nature of the surfaces remain unaltered, i.e. if F is the limiting value of the friction and N the normal reaction between two given surfaces, then the ratio $\frac{F}{N}$ is a constant. It is denoted by μ, and is termed the ***coefficient of limiting friction***, or the ***coefficient of static friction***. Hence

$$F = \mu N.$$

When the above slider is just about to move the forces acting on it are as shown in Fig. 3·18 (*b*), where R is the resultant of the normal reaction N and the friction F. The reaction R is inclined at an angle θ to the vertical, given by $\theta = \tan^{-1}\mu$. This angle is called the ***angle of friction***.

Experimental Determination of the Coefficient of Static Friction.—If the surface of the body under examination is flat, the coefficient of friction may be found as follows: The body is placed on a flat surface, AB, Fig. 3·19, pivoted about a horizontal joint at A. The other end, B, is attached to a bucket, C, into which lead shot may be poured to increase the tilt of the surface. Eventually a stage is reached when the body is just on the point of moving down the plane. Let θ be the inclination of the plane at this moment. The force acting down the plane is then $Mg \cdot \sin\theta$ which is equal and opposite to the frictional force, F, acting on the body. The normal reaction, N, the value for which is obtained by resolving forces in a direction normal to the plane, is $Mg\cos\theta$. We therefore have

$$\mu = \frac{F}{N} = \tan\theta.$$

The value of θ given by this equation is called the ***angle of repose***.

Kinetic Friction.—When slipping occurs between two bodies

in contact a frictional force continues to oppose the motion but, in general, the *magnitude* of this force is less than the frictional force existing just before slipping occurs. Experiment shows that as long as the motion is not too great, the frictional force F′ is directly proportional to the normal reaction between the surfaces and is independent of the velocity, i.e.

$$F' = \nu N$$

where ν is the coefficient of kinetic friction.

Suppose that a body of mass m rests on a horizontal table which is not smooth. Then $N = mg$, and $F' = \nu mg$ when the body is moving. Suppose F_1 is the force applied to the body. Since F_1 and F′ act in contrary senses, on a body of mass m, its acceleration a is given by

$$F_1 - F' = ma, \text{ or } a = \frac{F_1}{m} - \nu g.$$

In the absence of friction the acceleration would have been $\frac{F_1}{m}$, so that the effect of friction is to reduce the acceleration.

If the body is in motion and $F_1 < \nu mg$, a will be negative and the body will be brought to rest. To start the motion again a force greater than νmg will be required—it will be μmg.

Perry's Apparatus for determining the Coefficient of Kinetic Friction.—The essential parts of this apparatus are shown in Fig. 3·20 (*a*) and (*b*). A is a heavy wheel capable of rotation about a vertical axis,

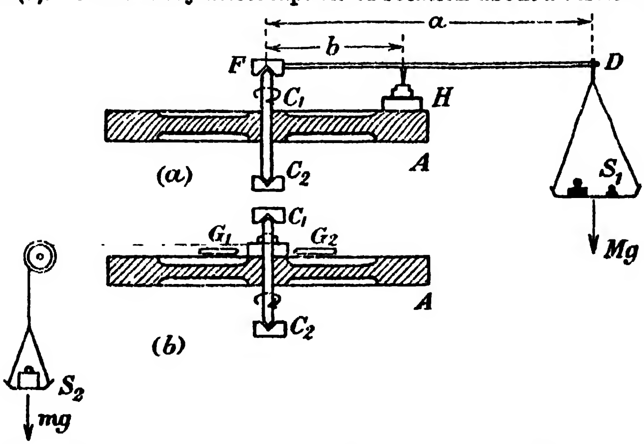


FIG. 3·20.—Perry's Apparatus for determining the Coefficient of Kinetic Friction between two Surfaces.

$C_1 C_2$. DF is a lever carrying a scale pan, S_1, and having its fulcrum on the vertical axis C_1C_2. When the pan is loaded, as indicated, a thrust is exerted by the lever on a block, H—the surface of this block in contact with the wheel is flat. Attached to this block is a pan, S_2—the attachment is made by a cord passing over a pulley. The coefficient

of kinetic friction to be determined is that appropriate to the surface of H and that of the rotating wheel. When the system is stationary, H rests against a stop, G_1. The wheel is rotated so that H remains about half-way between the stops G_1 and G_2—it is said to be in 'floating equilibrium.' This condition is obtained by varying the load in S_2. Under these conditions the friction is equal to mg, the weight of the pan S_2 and its load. The normal reaction between the surfaces in contact is $Mg \cdot \frac{a}{b}$, where M is the mass of S_1 and the load in it. [We neglect the mass of the lever.]

$$\therefore \nu = \text{coefficient of kinetic friction} = \frac{m}{M} \cdot \frac{b}{a}.$$

Example. A body of mass 4 lb., hanging freely over the edge of a rough table, is connected by means of a light string passing over a smooth pulley at the edge, to a body of mass 2 lb. resting on the table. This is pulled 2 ft. along the table in 0·5 sec. from rest. What is the coefficient of friction ?

Let F poundals be the friction ; T poundals the tension in the cord. Then the resultant force pulling the 2 lb. mass is T − F, so that its acceleration is given by

$$T - F = 2a.$$

Using $s = \frac{1}{2}at^2$, we have $a = 16$ ft. sec.$^{-2}$

$$\therefore T - F = 32 \text{ poundals.}$$

Considering the 4 lb. mass, the resultant downward force acting on it is

$$128 - T = 4 \times 16.$$

$$\therefore F = 32 \text{ poundals} = 1 \text{ lb.-wt.}$$

$$\therefore \nu = \frac{F}{2g} = 0·5.$$

The Friction Dynamometer.—The principle of this instrument, which is an application of the frictional forces existing between surfaces in contact to measure the rate at which work is done, is as follows: A large pulley wheel of radius r_1, Fig. 3·21, is rigidly fixed to the axle of the engine under test. A flexible belt having wooden blocks on its under side is placed over the outer rim of the wheel. One end of this belt carries a bucket W into which lead shot can be poured to increase its mass. The other end is fixed to a spiral spring attached to some rigid support [the floor]. Let r_2 be the outer radius of the belt. Suppose that the shaft makes n revolutions per second when the condition of 'floating equilibrium' has

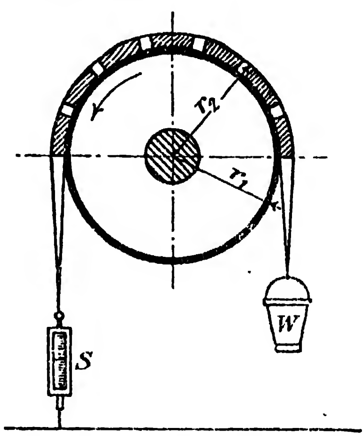


FIG. 3·21.—Friction Dynamometer.

[N.B.—The arrow on r_1 indicates that r_1 is the radius of the outer surface of the wheel.]

been obtained; let W be the weight of the bucket and its contents, while S is the reading on the spring balance. The moment about the axis of the shaft of the forces due to the weight W and the tension in the spring is $(W - S)\left(\frac{r_1 + r_2}{2}\right)$. This must be balanced by the moment of the frictional forces F about the same axis, viz. Fr_1. Now the distance through which the edge of the wheel moves against F is $2\pi r_1 n$ every second, so that the work done per second in overcoming friction is $2\pi n r_1 . F$. Eliminating F from this equation we find that the work done per second is $\pi n (r_1 + r_2)(W - S)$. This is the ***power*** of the engine, since it is the ***rate*** at which work is being done.

Rolling Friction.—To fix our ideas let us consider an engine wheel moving along a rail. There is never only contact at a point or even along a line normal to the rail, but always over a surface due to the elastic deformation of the bodies. Thus there is a small sliding motion of those parts of the surface in contact. There is thus brought into existence a frictional torque retarding the motion.

EXAMPLES III

1.—A force of 7 lb.-wt. acts on a mass of 29 lb. for 6 sec. How far has the body moved from rest? What is its final momentum?

2.—Find the resultant of two forces, 6 and 8 lb.-wt. respectively, acting on a body with an angle of $67\frac{1}{2}^\circ$ between them. If this resultant acts on a mass of 1 cwt., determine the acceleration. Construct the velocity-time curve for the first 4 sec. of its motion.

3.—Enunciate the theorem known as the parallelogram of forces, and describe an experimental arrangement whereby this law may be verified.

4.—A circular disc has a radius of 10 cm. At a point 7 cm. from its centre a circular hole 4 cm. in diameter is punched. Calculate the position of the centre of gravity of the remaining metal.

5.—A uniform beam 18 ft. long, whose mass is 1 cwt., is inclined at 60° to the vertical. It is held in position by means of a horizontal cord 13·8 ft. from its lower extremity. Calculate the tension in the cord.

6.—What force is required to raise a load of 2 cwt. by means of the second system of pulleys if there are 4 pulleys in the lower block? A similar load is also raised by means of Weston's differential pulley in which $R = 1$ ft. and $r = 11$ in. Compare the velocity ratios in the two systems.

7.—Describe a balance, indicating the features which a good balance should possess.

8.—A body requires 20·61 gm. to hold it in equilibrium when placed in one pan of a balance, and 20·73 gm. when placed in the other. Calculate its true mass.

9.—Two masses, 5 and 12 lb. respectively, are attached to the ends of a uniform rod 6 ft. long, mass 3 lb. Where must a 20-lb. mass be

placed so that the whole will balance about a point 2 ft. 6 in. from the 5-lb. mass?

10.—A mass of 10 gm. placed at the 97 cm. division on a uniform metre scale causes the whole to balance when the fulcrum is at the 55·3 cm. division. Calculate the mass of the scale.

11.—Two forces, 3·6 and 5·8 lb.-wt. respectively, have a resultant equal to 8·1 lb.-wt. What is the angle between the forces? Check by a graphical method.

12.—Derive an expression for the time of oscillation of a simple pendulum. Explain how the intensity of gravity may be determined by means of such a pendulum.

13.—A uniform board ABC in the form of an equilateral triangle of 12 in. side weighs 3 lb. and has weights of 4 lb. and 5 lb. hanging from A and B respectively. Find a point from which the board may be suspended so that it sets in a vertical plane with AB horizontal and C pointing down. Is there more than one such point? (L.S.C.)

14.—Explain, giving diagrams of the forces acting in each case, (*a*) how it is possible to sail a boat against the wind, (*b*) why the nose of a racing motor-boat rises out of the water, (*c*) why a railway ticket-collector leans backwards when alighting from a moving train.

15.—How would you compare accurately (*a*) the length of a standard yard with that of a standard metre, (*b*) the period of torsional oscillations of a horizontal rod suspended by a fine wire with that of a seconds pendulum?

16.—A uniform cylinder of height h and radius r rests with its plane base on a rough inclined plane. The angle of inclination of the plane may be increased gradually from zero. Show that the cylinder will topple over before it slides if $2r/h$ is less than the tangent of the angle of friction.

17.—What is the radius of the sharpest bend which may be turned without skidding by a motor-car travelling at 30 ml. hr.$^{-1}$ on a level road if the coefficient of friction is 0·7.

18.—A body slides from rest down a rough plane in 5 sec. If the coefficient of friction is 0·42, and the inclination of the plane 25°, what is the length of the plane?

19.—The distance between the scale-pan knife-edges in a balance is 30 cm. The central knife-edge is at a perpendicular distance of 1 cm. above the middle point of the line joining the scale-pan knife-edges. The centre of gravity of the beam is 2 cm. below the central knife-edge. The mass of the beam is 850 gm.; that of each scale-pan 100 gm. Find the deflection of the beam when masses of 50 and 51 gm. are placed in the pans.

20.—Explain the construction of a good beam-balance, pointing out the factors which determine (*a*) its accuracy, (*b*) its sensitiveness.

How could you find the mass of a body if you had to use a balance which was not true?

CHAPTER IV

THE ELEMENTS OF HYDROSTATICS

Density and Specific Gravity.—The density of a substance is defined as the mass of the substance per unit volume. *A priori* this statement calls for little comment, for whether 1 cm.^{3} or 1000 cm.^{3} are used in the experimental determination the same value for the density is obtained within the limits of experimental error. If, however, one adopts the modern view that all substances consist of molecules or atoms which are not in contact with one another, and which do not fill the whole of available space, some further remarks are necessary. Suppose that some imaginary being is free to move in and out amongst the molecules; his idea of the density of the medium will be very different from ours, for the particular volume which he chooses may contain many or a few such molecules, or even none at all. These statements are made here to show the student that some of our most commonplace ideas, i.e. ideas gained from a macroscopic view of things, are very different when the structure of matter is considered microscopically.

The idea of density is frequently confused with that of *specific gravity,* which is defined as the *ratio* of the mass of a given substance to that of an equal volume of water at the same temperature. Since this value is a ratio it is independent of the system of units used in the experimental determination, whereas the density, being a mass per unit volume, must always be expressed in gm. cm.$^{-3}$, or lb. ft.$^{-3}$, etc.

Fluids.—Solids are those substances which offer a considerable resistance to any force endeavouring to change their size or shape. On the other hand fluids, such as alcohol or nitrogen, cannot offer any permanent resistance to impressed forces tending to alter their shape. The term *fluid* is used to include both *liquids* and *gases,* the fundamental difference between liquids and gases being that the latter always occupy the whole of the space which is available, whereas liquids are always characterized by the presence of a free surface. This free surface is horizontal for such masses of liquid as are found in pools, etc., but becomes curved when the mass of

liquid is larger, as in the case of a sea ; in both instances the surface is everywhere perpendicular to the earth's radius at that point, but it is only in the second that the curvature can be detected easily. The thrust on any solid surface in contact with a fluid at rest is everywhere normal, i.e. perpendicular to the surface. If this were not so the thrust could be resolved into forces perpendicular and parallel to the surface, and this parallel force would cause motion of the body.

Pressure.—Whenever a force, F, is applied to an area, s, so that it is distributed equally and acts normally to the surface, then $\frac{F}{s}$, the force per unit area, is termed the ***pressure*** on that area. If the force is not distributed equally we may determine the pressure at any point on that area by constructing a small area round the said point. If δF is the force acting on such a

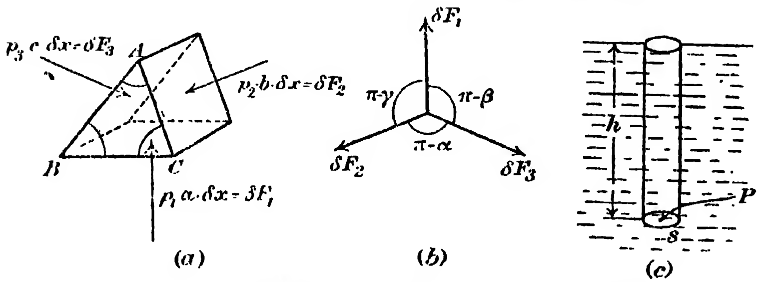


FIG. 4·1.—Pressure at a point in a Liquid at rest.

small area δs, then when δs is sufficiently small we may consider the force to be distributed uniformly over δs, so that the pressure is $\frac{\delta F}{\delta s}$; in the limit this becomes $\frac{dF}{ds}$.* [In the c.g.s. system the absolute unit of pressure is one dyne. cm.$^{-2}$; in the f.p.s. system it is one poundal. ft^{-2}. The corresponding gravitational units are the gm.-wt. cm.$^{-2}$, and the lb.-wt. ft.$^{-2}$]

To Show that the Pressure at a Point in a Fluid at rest is the same in all directions.—Consider any point in the fluid, and suppose that a wedge in the form of a triangular prism of arbitrary section ABC, Fig. 4·1 (a), surrounds the point. Then the fluid inside the wedge is in equilibrium under the action of

(i) its weight acting vertically downwards,
(ii) the thrusts on its faces.

If the wedge is very small the weight of the fluid in it, depending on the product of three small quantities, is negligible in comparison

* $\frac{dF}{ds} = \lim_{\delta s \to 0} \frac{\delta F}{\delta s}$.

with the forces acting on the sides, each of which depends on the product of two small quantities. Now the forces acting normally on the two ends of the prism are equal and opposite so that they may be omitted in the problem before us. Let the forces acting normally on the three other faces be δF_1, δF_2, and δF_3: these must be in equilibrium since the fluid is at rest. If the angles of the section are α, β, and γ, the angles between the lines of action of the forces are $(\pi - \alpha)$, $(\pi - \beta)$, and $(\pi - \gamma)$ respectively—cf. Fig. 4·1 (b). Then

$$\frac{\delta F_1}{\sin(\pi - \alpha)} = \frac{\delta F_2}{\sin(\pi - \beta)} = \frac{\delta F_3}{\sin(\pi - \gamma)}.$$

But $$\frac{\sin\alpha}{a} = \frac{\sin\beta}{b} = \frac{\sin\gamma}{c}, \qquad [a, b, \text{and } c \text{ are the sides of } \triangle ABC.]$$

$$\therefore \frac{\delta F_1}{a\,.\,\delta x} = \frac{\delta F_2}{b\,.\,\delta x} = \frac{\delta F_3}{c\,.\,\delta x},$$

where δx is the length of the wedge, i.e. the pressures over the faces of the prism are equal.

Pressure at a Point in a Fluid.—To determine the pressure at a point distant h below the free surface of a liquid which is at rest, and whose density is ρ, we imagine a small horizontal area s drawn round P and consider the liquid contained in the right cylinder having this area as its base—cf. Fig. 4·1 (c). This cylinder of liquid is in equilibrium under (a) the upthrust on the base, (b) its own weight, (c) the thrusts due to the pressure of the surrounding liquid on its sides. Since these are everywhere normal to the surface they have no vertical component, so that for equilibrium the weight, W, of the cylinder of liquid must be equal to the upthrust on the base.

Now, using the c.g.s. system of units,

W (dynes) = weight of a column of liquid of height h (cm.), and cross-sectional area s (cm.^{2}),
= mass of this column $\times g$, the acceleration due to gravity,
= [volume of this liquid, v, (cm.^{3}) $\times$ its density, ρ, (gm. cm.$^{-3}$)] $\times g$,
= $(sh\rho)g$ (dyne).

Hence, F, the total thrust on the area s is $(sh\rho)g$ (dyne). The pressure P at any point in the base is therefore given by

$$P = \frac{F}{s} = (g\rho h) \text{ (dyne. cm.}^{-2}).$$

From the above we see that the pressures at two points in the same horizontal plane in a liquid at rest must be equal. This may

be shown by cutting a piece of brass tubing at right angles to its axis and arranging the two new ends thus formed in the same horizontal plane. To facilitate this adjustment a flat sheet of metal and a spirit level may be used. A beaker containing liquid is then placed so that the ends of the brass tubes are immersed. When the liquid is at rest the ends of the tubes must be at the same depth below its surface. If the two tubes are connected together by means of a T-piece and rubber tubing, bubbles of gas appear from the two ends at the same time when pressure is applied to the open end of the T-piece. In carrying out this experiment narrow tubes must not be used since other forces become appreciable so that the simplicity of the experiment is lost: the reason for this will be noticed later [cf. p. 119].

Archimedes' Principle.—When a body is immersed either wholly or partly in a fluid at rest, it displaces a volume of fluid equal to that of the immersed portion, and experiences an upthrust due to the liquid displaced; the magnitude of this upthrust is equal to the weight of the displaced fluid. Let A, Fig. 4·2, be such a body. If the body is supposed to have been removed, and the space it occupied filled with some of the fluid, the forces arising from the superincumbent fluid are unaltered. Now the resultant of these forces just balances the gravitational force acting on this mass of the fluid, viz. its weight—the above resultant must act vertically upwards. When the body was in the fluid these forces were still existent and must therefore have reduced the effect of the earth's attraction on the body, i.e. its weight was apparently diminished by an amount equal to the weight of fluid displaced. If the body A were suspended from a balance this apparent loss in weight would be detected as an apparent loss in mass.

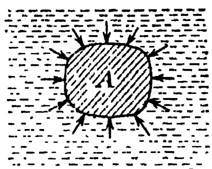


FIG. 4·2.

A similar argument holds when the body is only partly immersed in the fluid.

Experimental Verification of Archimedes' Principle.—The apparatus commonly employed to demonstrate the truth of the above principle is shown in Fig. 4·3. It consists of two cylinders A and B which are of such dimensions that the solid cylinder B just slides into A and fills it completely. When in this position the whole is suspended from the arm of a balance and the balance equilibrated, [sand may be used]. B is then withdrawn and suspended in a beaker containing liquid from below A with the aid of the hooks provided. The equilibrium of the balance is thereby destroyed, but it may be restored by pouring some of the same liquid into A as that in which B is immersed. Equilibrium

will be established when A is completely filled with liquid. This verifies that the upthrust on B when this is completely immersed in a liquid is equal to the weight of the liquid displaced by B.

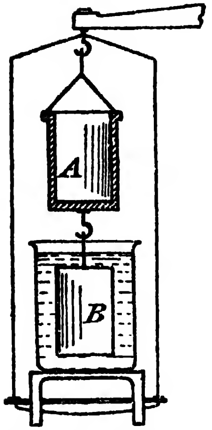


FIG. 4·3.— Apparatus to verify Archimedes' principle.

The experiment should also be repeated when B is a closed hollow cylinder such that it floats in the liquid. The procedure is exactly as above except that a piece of lead sufficient to cause the cylinder to sink is suspended from A and immersed in the liquid throughout the whole time that the experiment is being performed.

In order to vary this second part of the experiment the cylinder B is made of iron and mercury is the liquid used. A piece of tungsten, density 18·4 gm. cm.$^{-3}$, will be required to sink the iron. Brass or copper must not be employed in place of iron since these metals form amalgams with mercury.

The Principle of Flotation.—If the resultant of the forces acting on a body partly submerged in a liquid at rest and due to the liquid exactly balances the weight of the body, then that body floats. But it has been seen above, that the resultant of these forces is equal to the weight of the liquid displaced, so that for a floating body it may be said that the weight of the liquid displaced is equal to the weight of the body.

Experimental Methods for the Determination of Density.—(*a*) The density of a solid substance insoluble in water can be found by determining its mass in air and then in water. The apparent loss in mass of the solid is equal to the mass of the water displaced, and since 1 cm.^{3} of water has a mass of 1 gm. the volume of water displaced and hence the volume of the solid are known. For accurate work account must be taken of the fact that it is only when the temperature of the water is 4° C. and the external pressure 1 atmosphere that 1 cm.^{3} of water has a mass of 1 gm. If an experiment were made at 20° C. how could the volume of the solid be found? From tables it is known that the density of water at 20° C. is 0·998 gm. cm^{-3}. Suppose that the apparent loss in mass of the submerged body is 13·61 gm. Then the volume is equal to the mass divided by the density, viz.

$$\frac{13{\cdot}61}{0{\cdot}998} = 13{\cdot}63 \text{ cm.}^3$$

(*b*) If the body floats in water it must be caused to sink by using a heavy piece of metal called a ***sinker***. Let m_1 be the mass of the

body in air, m_2 the mass of the body in air plus the sinker in water, and m_3 the mass when both are suspended together in water. Now

$m_2 = m_1 +$ apparent mass of sinker in water,
 $= m_1 +$ mass of sinker in air $-$ mass of water displaced by sinker, and
$m_3 =$ apparent mass of both in water,
 $= m_1 -$ mass of water displaced by the solid $+$ mass of sinker in air $-$ mass of water displaced by sinker,
 $= m_2 -$ mass of water displaced by solid.
 $\therefore\ m_2 - m_3 =$ mass of water displaced by the solid.

If ρ_0 is the density of water at the temperature of the experiment, the volume of the water displaced is $(m_2 - m_3) \div \rho_0$: this is the volume of the solid. The density of the solid is therefore

$$\frac{m_1\rho_0}{(m_2 - m_3)}.$$

(*c*) The density or specific gravity bottle is a small glass container fitted with a ground glass stopper. A capillary hole in this stopper permits an excess of liquid to be removed and at the same time ensures a constant volume for the bottle. It is filled with the liquid and then cleaned and filled with distilled water, the mass of liquid in each instance being determined. The specific gravity of the fluid is the ratio of these masses; the density is easily calculated at any temperature as in (*a*). In using the bottle care must be taken to see that no air bubbles remain clinging to the sides of the bottle, and that the bottle has been completely filled at the same temperature in both instances.

(*d*) The density of a solid, available as a powder or as small crystals, may be determined with the aid of a density bottle. The method will be illustrated by considering how to determine the density of some crystals (e.g. sugar, copper sulphate, etc.) which are soluble in water but not in some other liquid (e.g. turpentine). The following observations must be made.

	Mass of bottle	$= m_1$
	,, ,, ,, + crystals	$= m_2$
	,, ,, ,, + crystals and turpentine to fill	$= m_3$
	,, ,, ,, + turpentine to fill	$= m_4$
	,, ,, ,, + water to fill	$= m_5$
	Mass of solid used . . .	$= (m_2 - m_1)$
	,, ,, turpentine required to fill bottle when crystals are present . . .	$= (m_3 - m_2)$
Now	,, ,, turpentine to fill bottle .	$= (m_4 - m_1)$
	,, ,, turpentine, the volume of which is equal to that of the crystals. . .	$= (m_4 - m_1) \quad (m_3 - m_2$

To find the volume of this mass of turpentine its density must be known. But the mass of water, of density ρ_0, required to fill the bottle is $(m_5 - m_1)$. The density of the turpentine is therefore

$$\frac{m_4 - m_1}{m_5 - m_1} \cdot \rho_0.$$

$$\therefore \text{ Volume of crystals} = \frac{(m_4 - m_1) - (m_3 - m_2)}{(m_4 - m_1)\rho_0} \cdot (m_5 - m_1).$$

$$\therefore \text{ Density of crystals} = \frac{(m_2 - m_1)\ (m_4 - m_1)\rho_0}{[(m_4 - m_1) - (m_3 - m_2)](m_5 - m_1)}.$$

(*e*) If the liquid whose density is required is only available in small quantities then its density may be found as follows:—A uniform glass capillary tube of suitable diameter (say 1 mm.) is selected, cleaned, dried, and its mass determined. A long length of mercury is placed in the tube, preferably by attaching a small piece of rubber to the tube, placing a bubble of mercury in the rubber and applying pressure at the open end of the rubber tube. This operation has a filtering action upon the mercury and enables the mercury to be introduced without undue contamination of the tube which is the result if suction is applied by the mouth. The mass of the mercury is determined. The tube is then filled with liquid and its mass found. In either case it is necessary to measure the length of the fluid in the tube. If very accurate results are required corrections to this length must be made owing to the existence of curved surfaces at the ends of the column. As a first approximation one adds (or subtracts) a length equal to two-thirds the diameter [1] of the tube, if the lengths have been measured as the distances between the extreme points at which the mercury (or liquid) is in contact with the glass. From the mass m, and corrected length l, of the mercury the mean radius of the tube is found, for if ρ is the density of the mercury at the temperature of the experiment, the volume of mercury is $\frac{m}{\rho}$ and this equals $\pi r^2 l$ so that

$$r = \sqrt{\frac{m}{\pi \rho l}}.$$

If r is small, m will also be small. It is then better to introduce in turn several pellets of mercury, measure the length of each, and determine their total mass, $\Sigma(m)$, say. Let this be μ—the only mass which has to be determined. If $\Sigma(l)$ is the total length of all the pellets,

$$r = \sqrt{\frac{\Sigma(m)}{\pi\rho\Sigma(l)}} = \sqrt{\frac{\mu}{\pi\rho\Sigma(l)}}.$$

[1] An approximate value of the diameter is obtained by finding a wire which will fit the tube and measuring its diameter with a screw gauge.

If M is the mass of a liquid whose density D is required, and this occupies a length L of the above tube, then

$$\frac{M}{D} = \pi r^2 L,$$

$$\text{or } D = \frac{M}{\pi r^2 L}, \text{ where } r \text{ is now known.}$$

In the above it was stated that the tube should be uniform in cross-section. This is only essential if the lengths of the mercury pellet and the column of liquid introduced are not equal, but a non-uniform tube may be used if the lengths of mercury and liquid columns are equal, for the tube may then be used as a density bottle of known volume.

The Internal Radii of Tubes.—The last paragraph has shown us how the internal radius of a narrow tube may be found, but the same method cannot be extended to wider tubes since the mercury would not fill the entire cross-section of the tube. We therefore proceed as follows :—A cork is inserted at one end of the tube and a little water (or mercury, if greater precision is desired) added so that when the tube is vertical the surface of the water is at some fiducial mark A. The mass of the whole is found. More water is then introduced until the level is at a second fiducial mark B. The mass is again determined and from the observations the volume of the tube between the marks A and B deduced. By proceeding in this way any errors due to the shape of the cork are avoided. If the length AB is known, the radius of the tube can be calculated. This same method may be used to find the radius of a test-tube. It should be noticed that this method, like the one above, only determines the *mean* radius of the tube.

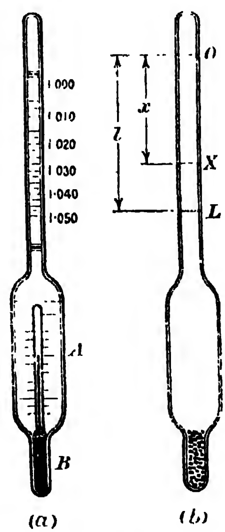


FIG. 4·4.

(a) A Common or Constant Mass Hydrometer.
(b) Theory of a Floating Hydrometer.

Hydrometers.—Two of the usual forms of hydrometer, which is an instrument used to determine the density of liquids or solids, are shown in Fig. 4·4; the first consists of a bulb A, at the lower extremity of which there is a small bulb B, containing mercury or lead shot. The neck between A and B is solid so that the mercury cannot be displaced. In the pattern shown here the bulb B is part of a mercury thermometer the scale of which is placed inside A. This enables the temperature of the liquid to be

observed without using a second thermometer. To the other extremity of A there is attached a long narrow tube which carries the scale of the instrument. The scale numbers generally refer to density, and the scale is so situated that the number at the point where the stem emerges from the fluid in which the hydrometer is immersed gives the density of the fluid.

The Equilibrium of a Floating Hydrometer (Elementary Theory).—Let us assume that the hydrometer is designed for use with liquids whose densities are greater than that of water. Let O, Fig. 4·4 (*b*), be the zero mark.

Let m = mass of hydrometer.

V = volume of the hydrometer up to its zero mark,

v = volume per unit length of the stem.

Then mg is the weight of the hydrometer, and this is the gravitational force pulling it downwards. Suppose that when the instrument floats in a liquid of density ρ the increase in length of the emergent part of the stem is n. Since the total volume of liquid displaced is $(V - nv)$, the upthrust of the liquid on the hydrometer is $(V - nv)\rho g$. For equilibrium

$$mg = (V - nv)\rho g,$$

i.e.

$$m = (V - nv)\rho.$$

The Graduation of a Common Hydrometer.—To calibrate this instrument, assuming that the stem is uniform in cross-section, one may proceed as follows. Suppose that O is the mark to which the instrument sinks when it is floated in water of density ρ_0 gm. cm.$^{-3}$; let L, Fig. 4·4 (*b*), be the mark when the hydrometer floats in a liquid of density ρ_1, this density being known or determinable. Let l be the distance OL. Let X be the mark on the stem to which the instrument sinks when floating in a liquid of density ρ. Call OX = x. The problem before us is to determine x in terms of l, ρ_1, and ρ: we then give values to ρ numerically equal to 1·00, 1·01, 1·02, etc., and so find out where these graduations must be placed.

If V is the volume of the instrument up to the mark O, and v the volume per unit length of the stem, we have, by the principle of flotation,

$$\begin{aligned} V \times \rho_0 &= \text{mass of water displaced} = \text{mass of hydrometer} \\ &= \text{mass of liquid displaced} \\ &= (V - lv) \,.\, \rho_1 = (V - xv)\rho. \end{aligned}$$

Hence

$$xv = V\left(1 - \frac{\rho_0}{\rho}\right), \text{ and } lv = V\left(1 - \frac{\rho_0}{\rho_1}\right),$$

so that

$$x = l\left[\frac{1 - \frac{\rho_0}{\rho}}{1 - \frac{\rho_0}{\rho_1}}\right].$$

In all accurate work with hydrometers it is very essential that the liquid surface should be clean. The following experiment verifies the above statement. A deep glass vessel is thoroughly cleaned and provided with a side tube near its base so that it may be completely filled with tap-water. A hydrometer is placed therein and the water allowed to overflow continuously. In this way a very clear water surface is obtained. The flow of water is stopped and the equilibrium position of the hydrometer noted. The water surface is then touched with a rod which has been wetted in a soap solution : this contaminates the water surface and the hydrometer rises—probably one or two millimetres. This is because the surface tension [cf. p. 113] of the liquid has been reduced and the hydrometer is not pulled down to the same extent as when the surface tension of the water had its maximum value, i.e. as when its surface was clean.

Nicholson's Hydrometer.—This instrument, which was designed for determining (*a*) the densities of solids and (*b*) those of liquids whose densities do not differ very much from that of water, consists of a hollow vessel, A, comprising a cylinder and two conical portions—cf. Fig. 4·5. The instrument carries upper and lower pans, B and C, respectively ; C is loaded with lead shot so that the hydrometer floats in an upright position when placed in a liquid. The hydrometer is made of brass and nickel-plated so that the tendency for air bubbles to cling to it shall be minimized. To find the density of a liquid the instrument is first placed in the liquid and masses added to the upper pan until a definite mark on the stem just touches the surface. It is generally somewhat difficult to judge this coincidence exactly so that it is better to solder a bent pin, P, to the stem of the hydrometer and always bring the point of the pin into contact with the surface of the liquid. This coincidence is best ascertained by looking at the reflexion of the pin in the surface of the liquid from a point below. If m_1 is the mass in the upper pan and M is the mass of the instrument itself, then, according to the principle of flotation, the mass

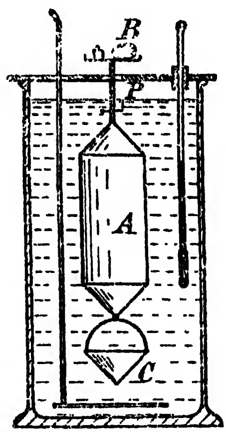


Fig. 4·5.—Nicholson's Hydrometer.

of the liquid displaced is $M + m_1$. The hydrometer is then washed and floated in water when a mass m_2 will be required in the upper pan in order to sink the instrument to P. The mass of the water displaced is $M + m_2$. If ρ_0 is the density of water at the temperature of the experiment the volume of water displaced is $(M + m_2)/\rho_0$. This is equal to the volume of liquid displaced. The density of the liquid is

$$\frac{(M + m_1)\rho_0}{(M + m_2)}.$$

If the instrument is floated in a liquid whose density differs considerably from unity, there is a tendency for it to tilt. This may be avoided by placing a suitable piece of brass in the lower pan during this part of the experiment, and making a correction as follows. Let m_1 = mass in the upper pan required to sink the instrument to the mark P when the piece of brass of mass μ and density ρ_1 is placed in the lower pan. Then the mass of the liquid displaced by the hydrometer is $M + m_1 + \mu$, and the volume of the liquid displaced, being the volume of liquid displaced by the hydrometer alone plus the volume of the piece of brass, is $\left(\frac{M + m_2}{\rho_0}\right) + \frac{\mu}{\rho_1}$. The density required is therefore

$$\frac{(M + m_1 + \mu)}{\left[\frac{M + m_2}{\rho_0} + \frac{\mu}{\rho_1}\right]}.$$

If a hydrometer is properly used, reliable results are obtained even for liquids whose densities do not differ much from unity,

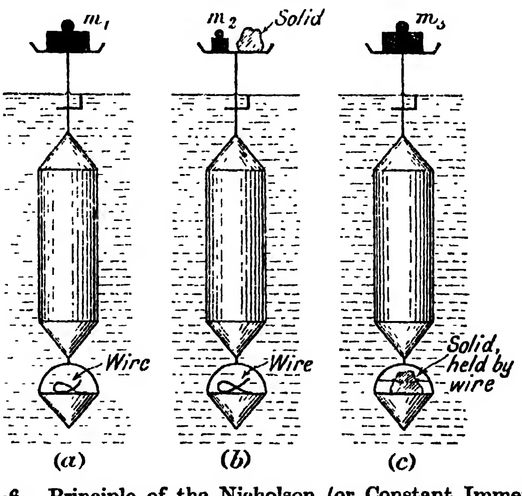


FIG. 4·6.—Principle of the Nicholson (or Constant Immersion) Hydrometer.

since the masses of large volumes of liquid and water have to be determined.

To determine the density of a solid the hydrometer is first floated in water as before and the mass required to sink the hydrometer to P ascertained. Let this be m_1, cf. Fig. 4·6 (*a*). The solid is then placed on the upper pan and the mass necessary to sink the instrument to the same mark again found, cf. Fig. 4·6 (*b*). Let this be m_2, so that the mass of the solid in air is $(m_1 - m_2)$. The solid is then placed on the lower pan when it will be found that a mass m_3 is necessary to sink the hydrometer to the same fiducial mark, cf. Fig. 4·6 (*c*). This mass will be greater than m_2 due to the upthrust of the water on the solid. Now by the principle of flotation, the mass of the water displaced in each instance is equal to the mass of the floating object. Hence, considering the state of affairs indicated in Fig. 4·6 (*c*), we have

Mass of water displaced by hydrometer when floating as in (*a*) + mass of water displaced by solid

$$= M + m_3 + \text{mass of solid in air.}$$

$\therefore$ $(M + m_1)$ + mass of water displaced by solid

$$= M + m_3 + m_1 - m_2.$$

$\therefore$ Mass of water displaced by solid

$$= (m_3 - m_2).$$

If the density of water is ρ_0, the volume of the solid is $\left(\dfrac{m_3 - m_2}{\rho_0}\right)$, so that its density is $\left(\dfrac{m_1 - m_2}{m_3 - m_2}\right)\rho_0$.

It will be noticed that this method applies equally well to solids which float, the only difference being that the solid must be tied to the lower pan. This may be done with the aid of a piece of wire and if this is allowed to remain on the lower pan throughout the experiment its mass need not be known.

Alcoholometry.—The term alcoholometry is applied to the determination of the strength of spirits. In the days of the alchemists rough-and-ready means were used. A piece of cloth was moistened with the spirit and a light applied: ignition indicated strong spirit. Sometimes an oil was poured upon the surface of the spirit; strong spirit floated on the surface of the oil. Later the spirit to be tested was used to moisten gunpowder—when a light was applied rapid combustion indicated a strong spirit; steady burning indicated a spirit which was regarded as 'good, rightfull and of vertue' and was known as 'proof' spirit. In 1666 some friction arose between importers of French brandy and the customs officials concerning the rate of duty chargeable on the liquid. There were two rates, 4*d*. and 8*d*. per gallon, for liquors of different qualities, and the revenue officials, guided by the sense of taste, asked for the higher rate. The decision was contested by the importers, but was eventually ratified; the

test was made statutory in 1670. Fraudulent merchants, however, attempted to disguise the taste of their brandies, and so other means had to be found. BOYLE first thought of using a hydrometer for testing spirits, and after various improvements it has become the standard instrument for such purposes.

'Over' and 'Under' Proof.—The term 'proof' is applied to spirits having a density 0·91976 gm. cm.$^{-3}$ at 15·56° C. (60° F.); this corresponds to 49·28 per cent. of alcohol by weight or 57·10 per cent. alcohol by volume. If the *over*-proof strength is added to 100, the sum represents the number of volumes of spirit at proof strength which that particular over-proof strength would make. Thus, 100 vol. of spirit at 16° over-proof are equivalent to 116 vol. of proof spirit, whereas 100 vols. of 16° under-proof are equivalent to 84 vol. of proof spirit. Absolute alcohol is 75·35° over-proof.

Sike's Hydrometer.—This is the particular form of instrument used in alcoholometry. It consists of a gilded brass bulb, 1·5 in. in diameter, to the bottom of which is fixed a counterpoise. The stem is a thin rectangular strip graduated in arbitrary units. Tables are supplied which convert readings into terms of over- or under-proof strengths.

Stability of Floating Bodies.—The principle of flotation [cf. p. 74] asserts that the mass of the floating object is equal to the

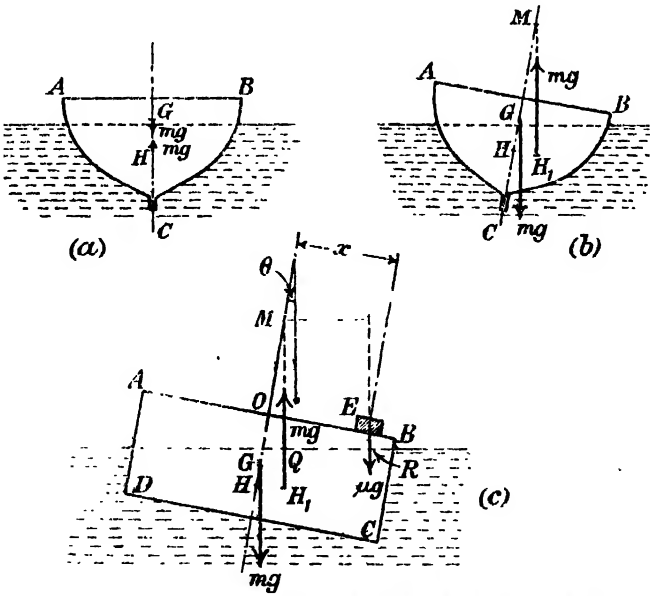


FIG. 4·7.—The Metacentric Height of a Floating Body.

mass of the liquid displaced. This condition alone is not sufficient to determine the equilibrium of the floating object. If it is in

equilibrium the weight of the solid must not only be equal to the upthrust of the liquid displaced, but these two forces must act in the same straight line.

Now while these two conditions are sufficient to determine the equilibrium of the floating body, the stability of that equilibrium requires further discussion.

Consider a floating body (e.g. a ship) in the position of equilibrium. If m is the mass of the ship, it is in equilibrium under the action of two forces, its weight, mg, where g is the acceleration due to gravity, acting vertically downwards through G, Fig. 4·7 (a), the centre of gravity of the body, and the upthrust, also mg, acting vertically upwards through H, the centre of gravity of the displaced liquid. The point H is termed the centre of buoyancy. Now HG is vertical when the ship is in its equilibrium position. We shall assume that this line is marked on the ship and that it moves with it when the equilibrium is disturbed. When the ship is displaced through a small angle, let the centre of buoyancy move to a position H_1, Fig. 4·7 (b), in the plane of the diagram. The mass of displaced liquid will remain unaltered, but its resultant upthrust will now act vertically through H_1. If this line of action of the upthrust cuts HG produced in M, then M is the ***metacentre*** of the ship, while the distance GM is the ***metacentric height*** of the ship.

The ship is now acted upon by a couple and if M is above G this couple will tend to restore the ship to its equilibrium position, i.e. the equilibrium is stable. Unstable equilibrium follows when M is below G.

To Determine Experimentally the Metacentric Height of a Rectangular Piece of Wood Floating in Water.—Consider that rectangular section ABCD, Fig. 4·7 (c), of the floating body which passes through G, the centre of gravity of the body. Suppose that the body is displaced through a small angle θ by placing a body of mass μ at E. Let M be the metacentre whose position is to be determined experimentally, and suppose that GM cuts AB in O. Let OE $= x$. If m is the mass of the wood, and μ is small compared with m, so that the mass of the displaced liquid may be considered constant, the three forces maintaining the body in equilibrium are its weight mg, acting vertically downwards through G, the upthrust mg acting vertically upwards at H_1, the centre of buoyancy in the disturbed position of the wood, and the weight μg of the mass at E which acts vertically downwards. By taking moments of forces about Q, the point of intersection of the water line with H_1M, we obtain GM, for

$$mg \,.\, \mathrm{GM} \,.\, \sin\theta = \mu g \,.\, \mathrm{RQ},$$

where R is the projection of E on the water line. If θ is small, $\sin\theta = \theta$, and $RQ = x\cos\theta = x$.

Hence $$GM = \frac{\mu x}{m\theta}.$$

The angle θ is deduced from observations on the position of a plumb-line attached to the wood as indicated.

Pressure of the Atmosphere.—The earth is surrounded by an envelope of mixed gases consisting of oxygen and nitrogen for the main part, but also containing carbon-dioxide, water vapour, and in smaller amounts argon, neon, krypton and xenon. This mixture is a fluid and, as such, exerts a pressure. In general, this pressure diminishes with increasing altitude, and is such that at distances greater than 50 miles above the earth's surface, the air is so rarefied as to be almost non-existent. Fig. 4·8 shows the effect of placing a tube, completely filled with mercury, in a reservoir of this substance. Whether the tube is inclined or not, the vertical height of the column, providing the mercury does not fill the tube entirely, is the same in each tube and is a measure of the pressure of the atmosphere under the prevailing conditions. The vacuum above the upper surface of the mercury is called a ***Torricellian*** vacuum, and should contain only traces of mercury vapour. This space is so called because it was discovered in 1643 by an Italian named TORRICELLI. Such tubes are the essential part of all mercury barometers.

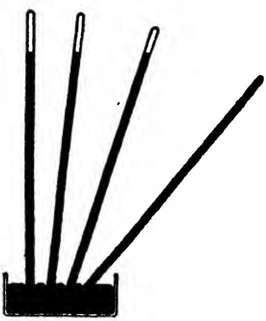


FIG. 4·8.

The Fortin Barometer.—The distinctive feature of this instrument, Fig. 4·9, is the device used for keeping the level of the mercury in the reservoir constant. This permits the use of a fixed scale—generally engraved on the brass case, A, surrounding the barometer tube, B. The reservoir bottom, C, is made of chamois leather and is moved by means of a plunger, the motion being imparted by the rotation of the screw S; this is moved so that the mercury level in the reservoir is coincident with the extremity of an ivory point P, whenever observations are being made. The tip of P coincides with the zero of the scale on A. The above coincidence is examined by viewing the reflexion of the point in the mercury surface. To determine the position of the upper surface of the mercury on the scale of A, the tube E, sliding inside A and operated by the milled knob D, is adjusted so that its lower end is level with the mercury surface. A vernier scale on E enables the position of the mercury

surface to be determined. After some months' use air tends to find its way along the glass-mercury surface; this is prevented from reaching the vacuum by means of the re-entrant glass joint X. The glass tube used in such a barometer is shown in Fig. 4·9 (*b*).

Boyle's Law.—Gases are fundamentally different from solids and liquids. The fact that a given mass of gas is at a certain temperature does not define its volume definitely, for a gas always occupies the whole of the available space in the vessel enclosing it. If the volume of the gas is increased the gas still fills the whole of the vessel, but the pressure it exerts on its walls is reduced. Similarly, if the volume is decreased, the pressure is increased. BOYLE, in 1662, investigated the relationship between the volume of a given mass of gas and the pressure to which it is subjected, and his results are expressed by the law which bears his name: '***The volume of a given mass of gas at constant temperature is inversely proportional to the pressure to which it is subjected.***'

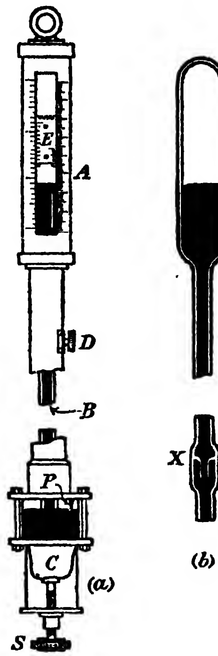


FIG. 4·9.—A Fortin Barometer.

Experimental Verification of Boyle's Law.—Fig. 4·10 (*a*) is a diagrammatic representation of the essential parts of the apparatus. It consists of a burette or other suitably calibrated vessel, A, connected by means of thick rubber tubing to a wide tube, B, containing mercury. C is a two-way tap leading either to a tube D, containing calcium chloride, or to a tube E. At the top of D there is a rubber bung through which pass E and another tube F which may be closed by a small glass cap and piece of rubber tubing. A loosely packed plug of glass wool, G, at the lower end of D prevents particles of the chloride from entering A. The tap C is first placed so that connection is made between A and E [cf. Fig. 4·10 (*c*)]. When B is raised the air or other gas in A is expelled into D via the tube E. During this operation care should be exercised to prevent the mercury from coming into contact with the grease on the tap, for mercury is easily contaminated. C is then rotated so that there is direct connection between D and A [cf. Fig. 4·10 (*b*)]. When B is lowered dry gas enters A. This operation is repeated several times so that the gas finally left in A is dry. With the tap C closed, B is raised to a considerable

height. If the mercury level in A continues to change, the tap C is leaking, so that this defect must be remedied before proceeding.

When it has been shown that the apparatus is free from leaks the volume of gas in A is noted, and the levels of the mercury in A and B are observed by means of the scale S. The difference between these two observations is a measure of the pressure difference between that in A and that of the atmosphere. If the barometric height is observed, the pressure of the gas in A in terms of cm. of mercury at room temperature may be deduced. A series of observations with the pressure in A both greater and then less than atmospheric is made. If now a graph is drawn showing the relation between p, the pressure, and V, the volume of the gas, a curve is obtained, but its nature cannot be directly inferred. But since it is expected that the observations will support the relationship $p \propto \frac{1}{V}$, i.e. $pV = \kappa$, where κ is a constant, we should plot $\log p$ and $\log V$. If the points lie on a straight line whose slope is -1, the validity of Boyle's law over the range of pressures investigated will have been established, for $\log p + \log V = \log \kappa$ = constant, is the equation to a straight line whose slope is -1. The validity may also be tested by plotting p against $\frac{1}{V}$, when a straight line should be obtained. Its slope is κ.

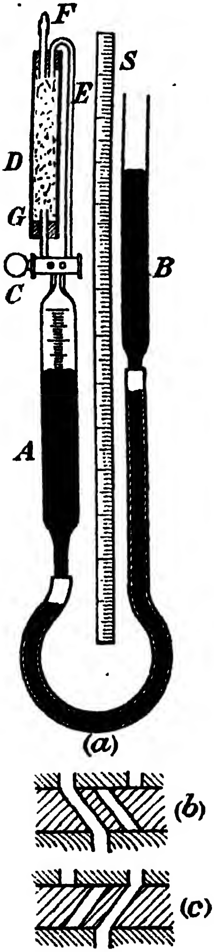


FIG. 4·10.—Boyle's Law Apparatus.

The actual method used by Boyle (1662) to establish his law for air was to observe the volume of air in the closed limb of a U-tube at atmospheric pressure and then at different pressures. He assumed the law to be valid and calculated what the volume should be for the pressures applied. This calculated volume was compared with the observed volume and the agreement was found to be very good.

[The numbers in Ex. 26, p. 101, have been taken from Boyle's original paper.]

Experiment. Clean and dry a glass tube about 40 cm. long and 0·3 cm. in diameter. Introduce a pellet of mercury about 10 cm. long into the tube. Observe the barometric height. Determine the

length of the tube occupied by the enclosed air when the tube is vertical and also when the tube is rotated through 180° in a vertical plane, i.e. when the pressure of the enclosed gas is greater and then less than atmospheric by an amount depending on the length of the mercury pellet. Introduce other pellets into the tube and repeat the observations. Hence investigate the validity of Boyle's law.

Hare's Density Apparatus.—This apparatus enables us to compare the densities of two liquids, so that if the density of one is known, that of the other may be deduced. It consists of two vertical tubes, AB and CD, Fig. 4·11. The upper ends of these tubes are connected to a T-piece and stop-cock, E: their lower ends each dip into one of the liquids under examination. By applying suction at E the liquids may be brought to convenient positions in the tubes. Let us suppose that these positions are P_1 and Q_1 respectively. If D and d are the densities of the two liquids while H_1 and h_1 are equal to the heights of P_1 and Q_1 above the exposed surfaces of the liquids, the difference in pressure between the inside and outside of the apparatus is $g\mathrm{D}\mathrm{H}_1$ or gdh_1, i.e. $\frac{d}{\mathrm{D}}=\frac{\mathrm{H}_1}{h_1}$. In actual practice it is at least inconvenient, and certainly undesirable, to adjust the ends of the scales S and T so that they are in contact with the exposed surfaces of the liquids. To avoid this, a long pin (or screw) is pushed through a piece of wood resting on top of the containing vessel in each instance, the pins being vertical, and their positions adjusted until their lower ends just touch the liquid surfaces. S and T are then used to measure the heights of P and Q above the tops of the pins, and if the lengths of the pins are known, H_1 and h_1 are easily deduced.

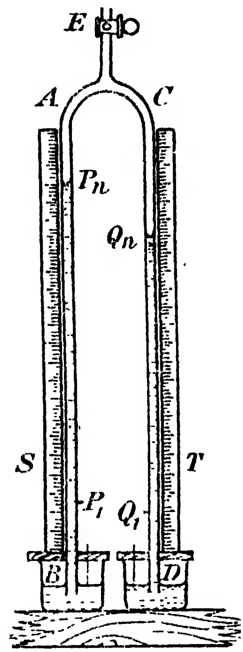


FIG. 4·11.—Hare's Apparatus.

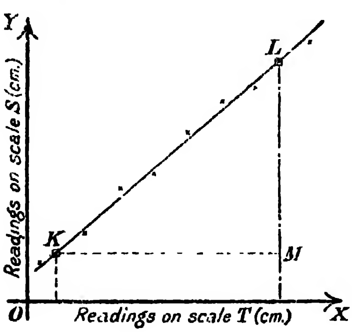


FIG. 4·12.

A series of observations with the levels of the liquids at different positions in the tubes is made, care being taken to see

that the tubes are thoroughly wetted. The observations are then plotted as in Fig. 4·12 and the best straight line drawn. Let K and L be two points on this line and draw KM and LM parallel to the axes of reference. It follows that ML and KM will be proportional to the same change of pressure inside the apparatus, so that if we denote them by H and h respectively, $g\mathrm{DH} = gdh$, i.e. $\frac{d}{\mathrm{D}} = \frac{\mathrm{H}}{h}$. Generally the liquid in AB is water so that in the c.g.s. system of units $\mathrm{D} = 1$ gm. cm.$^{-3}$, and therefore $d = \frac{\mathrm{H}}{h}$ gm. cm^{-3}

Buoyancy in Gases.—It has already been shown that any solid immersed in a liquid experiences an upthrust equal to the weight of the liquid displaced. Gases, too, exert an upthrust on bodies in them equal to the weight of the gas displaced. This may be demonstrated in the following manner. A, Fig. 4·13, is a

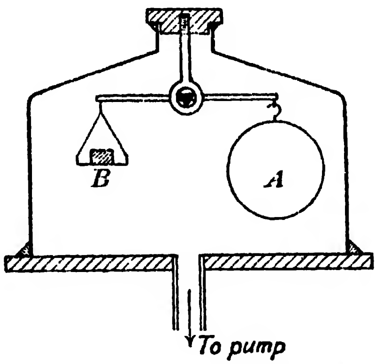


FIG. 4·13.—Buoyancy in Gases.

hermetically sealed vessel—a glass globe, for example—suspended from one arm of a balance and counterpoised by a mass, B. The whole is placed inside a large bell-jar which may be exhausted. As the air is removed from the jar the up-thrust on the large body A is much reduced in comparison with that on the counterpoise B. In consequence, the equilibrium of the balance is destroyed and A falls.

Correction for Buoyancy in Determining the Mass of an Object.—Let m be the mass of the weights (brass) necessary to counterpoise a given object, the weighing operation being carried out in air. Let ρ_1 be the density of brass, ρ_2 that of the material of the solid whose mass is being determined, and ρ_a that of the air under existing conditions. Let M be the true mass of the solid.

Then its volume is M/ρ_2, so that the upthrust on it due to the air displaced is

$$\left(\frac{M}{\rho_2}\right)\rho_a \cdot g.$$

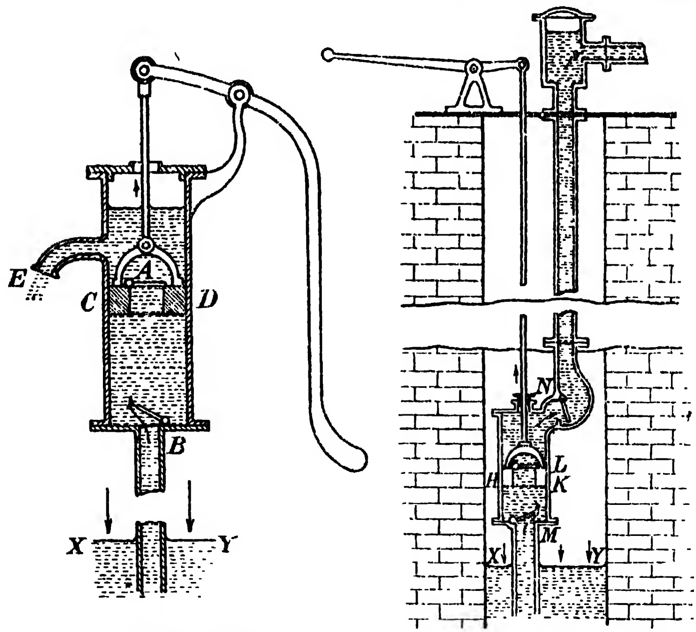
On the brass weights the upthrust is $\left(\frac{m}{\rho_1}\right)\rho_a \cdot g$. For equilibrium

$$Mg - \left(\frac{M}{\rho_2}\right)\rho_a g = mg - \left(\frac{m}{\rho_1}\right)\rho_a g$$

i.e.
$$M\left[1 - \frac{\rho_a}{\rho_2}\right] = m\left[1 - \frac{\rho_a}{\rho_1}\right]$$

$$\therefore\ M = m\left[1 - \rho_a\left(\frac{1}{\rho_1} - \frac{1}{\rho_2}\right)\right], \text{ since } \rho_a \text{ is small.}$$

The Suction Pump.—A diagrammatic representation of the suction or bucket pump is shown in Fig. 4·14 (*a*). The valves A and B are so constructed that they can only move upwards;



(*a*) A Suction Pump. (*b*) A Lift or Force Pump.

FIG. 4·14.

when the piston or bucket CD is forced downwards any water between the valves A and B is compelled to pass upwards through A, for the valve B is closed. When the motion of CD is reversed, i.e. the piston moves upwards, the water above it closes

the valve A and this water is carried upwards and delivered through the spout E. The space between CD and B would now be a vacuum were it not for the fact that the atmospheric pressure acting on the surface XY of the water in the reservoir forces the water past the valve B into the cylinder of the pump. On the descent of CD the cycle is repeated, the result being an intermittent delivery of water from the pump. In dry weather it is often necessary to *prime* such pumps, i.e. water must be poured into the main body of the pump in order to make an air-tight seal at CD. If such a process is not used the pump will not work.

The Lift Pump.—As in the preceding pump, there are two valves L and M, Fig. 4·14 (*b*), and an additional valve N is in a side exit. When the piston HK is raised the valve L closes, while M and N open, allowing water to pass from the reservoir into the cylinder below HK, while the water above HK is forced through N upwards into the cylinder. On the downstroke of the plunger HK the valves M and N close, and the water is forced through L into the receptacles which are being fed. The cycle of operations is then repeated. Vessels are often fitted to plunger pumps in order to provide a 'cushion' and so avoid damaging the pump when the piston motion is reversed. The air cushion absorbs the shocks which are due to the alternate starting and stopping of the water supply.

The Limitations of the Above Pumps.—Under normal conditions the pressure of the atmosphere is sufficient to support a column of mercury 30 in. in length. Since mercury has a density 13·6 times that of water the height of a water column which can be supported under similar conditions is $30 \times 13{\cdot}6$ in. or 34 ft. This distance represents the maximum theoretical distance between the water-level XY and the valve B, cf. Fig. 4·14 (*a*). In practice, owing to imperfections in the pump, it is seldom found that water can be raised more than 20 ft. by a suction pump.

This distance must not be confused with the height to which water can be driven by means of the force pump. This latter height depends upon the efficiency of the pump and the strength of the valves. A distance of 300 ft. is about the maximum distance through which it is safe to raise water in this way.

The Petrol Pump.—The lift pump finds a useful application in the modern petrol pump for raising petrol from an underground tank. When the plunger is raised by the ratchet work, R (shown in the conventional manner), Fig. 4·15, the valves V in the piston are closed and W is opened so that the petrol rises; on the descent of the plunger W automatically closes, thereby preventing the petrol from flowing back into the tank. At the same time the valves V are opened and the petrol is forced upwards into the glass vessel A, the air in A escaping through the outlet C. When A is filled, any excess of petrol

driven into it by the lift pump escapes down B and returns to the tank. The petrol in A is delivered through the tap T.

The Siphon.—The siphon, Fig. 4·16 (*a*), consists of a piece of tubing of rather small bore (0·5 cm.) bent so that its two arms are unequal. If the tube is filled completely with liquid and the shorter arm is immersed in a liquid, liquid is removed from the containing vessel. The column of liquid BC exerts a pressure at C, and when the siphon begins to operate the liquid runs out at C. The removal of the liquid from this side of the siphon tends to produce a vacuum in BA, and consequently the liquid is drawn from the reservoir, which is being emptied, into the tube. The whole process becomes continuous so that there is a steady stream of liquid at C. The speed at which the liquid is removed from its container depends upon the vertical distance between the level of the liquid and C; the greater the distance, the more rapid the flow from the siphon. However, it must be noted that if the vertical distance between A and B exceeds the barometric height, expressed in terms of the liquid in A, then the column AB can no longer be maintained and the siphon ceases to work. For water the above distance is 30 ft. (about), for mercury, 76 cm. The above argument indicates that a siphon will not work in a vacuum.

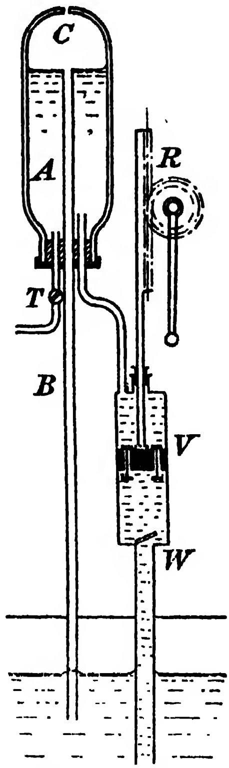
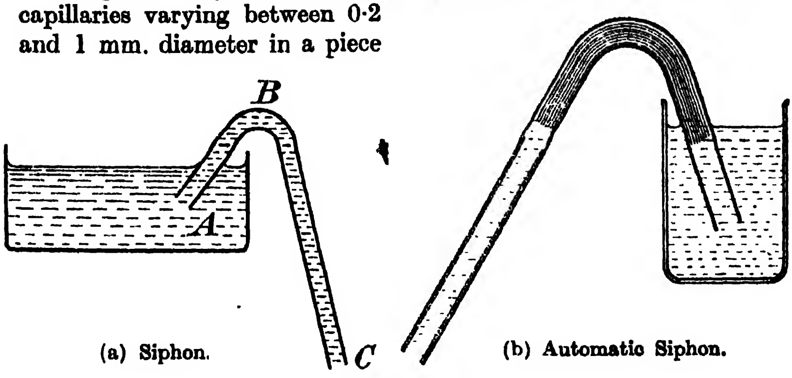


FIG. 4·15.—A Petrol Pump.

A siphon may be rendered automatic by placing some capillaries varying between 0·2 and 1 mm. diameter in a piece



(a) Siphon. (b) Automatic Siphon.

FIG. 4·16.

of straight glass tubing and bending the whole so that the shape shown in the Fig. 4·16 (*b*) is obtained. A little molten wax, made by melting together 10 parts resin and 6 parts vaseline, is drawn into the longer limb of the siphon so that the walls of the glass are thinly coated. When the shorter limb is placed in a liquid, capillary action causes some to pass into the waxed limb and form a pellet. This grows until the vertical distance between its ends exceeds the depth of the end of the short limb below the liquid surface. The ordinary action of a siphon ensues.

The Hydraulic Press.—A modern form of the hydraulic press first invented by BRAMAH is shown in Fig. 4·17. It consists essentially of a large cylinder, A, filled with water (or oil) in communi-

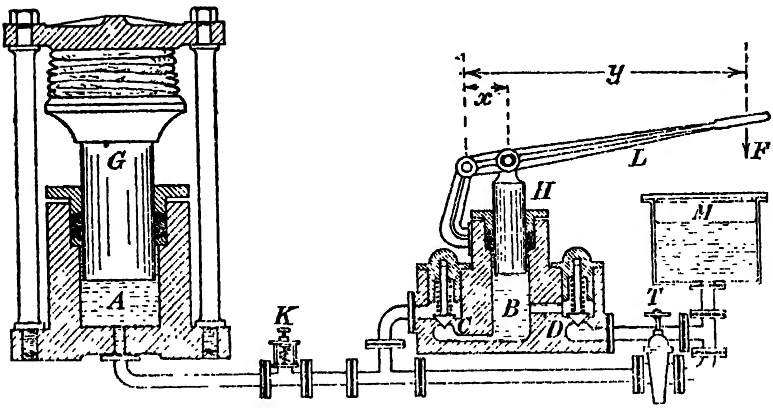


FIG. 4·17.—Hydraulic or Bramah Press.

cation with a smaller one, B. The larger cylinder is provided with a piston, G, known as the press-plunger, while the smaller one is provided with a piston, H, of much less cross-sectional area. It is termed the pump-plunger. Packing glands prevent the escape of liquid from the junctions between the pistons and the respective cylinders. H is operated by means of a lever, L, which further increases the mechanical advantage of the press. When a thrust is applied to the top of the smaller piston the pressure in B increases so that a valve, C, opens and the pressure is transmitted to the liquid in A. In consequence of this the press-plunger rises and compresses any goods carried on a platform attached to the top of G. When the lever is raised the valve C closes and D opens so that liquid enters B. The process may then be repeated. If, through some defect, the piston G fails to respond to the increased force acting upon it, the safety-valve K opens and the escaping liquid returns to the reservoir M via a channel not indicated in the diagram.

To release the pressure on the liquid in A the tap T is opened and the liquid returns to M.

If x and y are the perpendicular distances from the fulcrum of the lines of action of the thrust on the smaller piston and of the effort F applied to the extremity of the lever, the thrust on H is $F\left(\frac{y}{x}\right)$. If s is the area of cross-section of H, the pressure on the liquid in B is

$$\frac{F\left(\frac{y}{x}\right)}{s} = \frac{Fy}{xs}.$$

If S is the cross-sectional area of G, the thrust on its base is

$$F \cdot \frac{y}{x} \cdot \frac{S}{s}.$$

The mechanical advantage of this machine is $\frac{y}{x} \cdot \frac{S}{s}$, i.e. it is the product of the mechanical advantage of the lever and that of the simple press. [The machine is here considered to be an ideal one.]

It must be noticed that in the above argument we have assumed that the pressure on the base of H is exactly the same as that on the base of G. This is only true when these are in the same horizontal plane. If, at any instant, h is the difference in the above levels, the pressure difference is $g\rho h$, where g and ρ have their usual significance. The correction to be applied to obtain the pressure on the base of G is therefore variable; in general it is positive at the beginning of the stroke and negative at the end of it.

Air Pumps.—The simplest form of air pump is the glass filter pump shown in Fig. 4·18. The tube A is connected to the water supply, while the side tube C leads to the apparatus to be exhausted. A rapid stream of water is forced along A, and this produces a jet of water which passes down the tube B. The air in the immediate vicinity of B becomes entrapped in the water stream and is carried away through D. This process of entrapping the air is continuous until a pressure of about 3 cm. of mercury is reached—the pump then ceases to reduce the pressure further.

If a lower vacuum is required some other form of pump must be employed; if the space to be exhausted is not greater than 200 cm.^{3} the modified Toepler pump, Fig. 4·19, is very useful. It consists of a cylindrical barrel A, about 200 cm.^{3} capacity. At its upper end is a two-way capillary tap T; by turning this tap the barrel A can be put into connection, either with the tube B, which leads to the apparatus to be exhausted, or with C, which is open to the air. At the lower end of A is a smaller barrel D, with a side tap attached; any air entering the apparatus via

the pressure tubing is entrapped in D and can be removed through this side tap. D is connected to a mercury reservoir E, by means of pressure tubing.

To commence operations the reservoir E is raised, T being connected to C, so that the mercury fills the barrel A completely. T is closed; E is then lowered a little and T rotated so that B and A are in connection. The pressure of the gas in B and the vessel to which it is attached forces the mercury downwards in A; E is

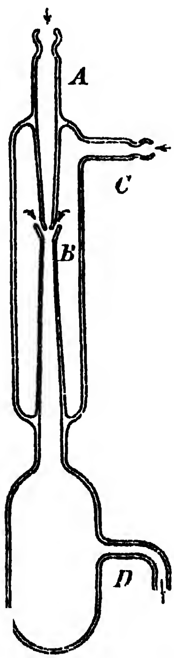


Fig. 4·18.—A Filter Pump.

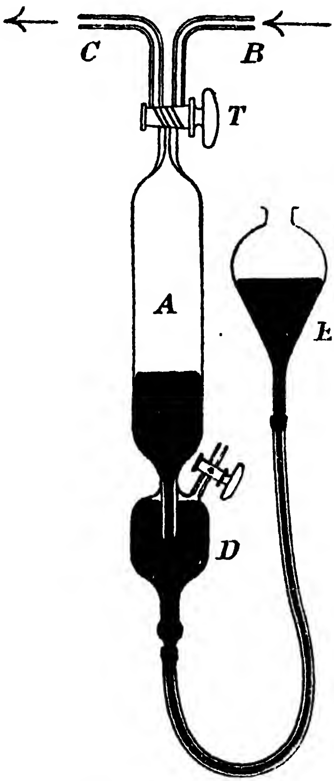


Fig. 4·19.—Toepler Vacuum Pump.

lowered until A is nearly filled with the gas. T is then closed and E raised until the pressure in A is greater than atmospheric. When this is so, T is put into connection with A and C so that the gas can be removed from A. The operation is repeated ten times or more, after which it will be found that no more gas can be removed from the vessel which is being exhausted. When the mercury in A reaches the tap T, the sound of a good metallic click indicates that a low vacuum has been reached.

The Sprengel Pump.—A form of this pump working in conjunction with a water pump is shown in Fig. 4·20. The capillary tubes in it are 0·15 cm. in diameter, the others about 0·5 cm. except where they widen out into bulbs approximately 2 cm. in diameter. The tube A leads to the vessel being exhausted. Pellets of mercury fall from the jet B and entrain bubbles of gas as they enter the fall tube below. The supply of mercury in B is replenished from the reservoir E which is in direct communication with a water pump. A capillary tube passes down the centre of this reservoir, through its base, and ends in the trough C. At the end of this tube there is a T-piece to which is attached a fine-drawn-out glass tube by means of a stout rubber tube. When the water pump is operating air is drawn in through this orifice and carries bubbles of mercury with it. When this mixture arrives at the upper end of the tube the air passes to the water pump while the mercury falls into the reservoir. A clip, K, controls the rate at which air enters the apparatus.

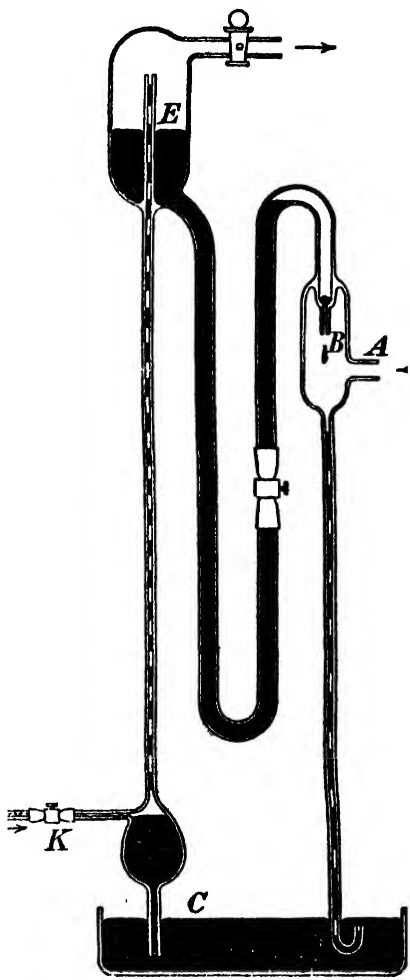


FIG. 4·20.—A Sprengel Pump.

High Vacua.—When the above procedure has been duly carried out, the degree of vacuum may be increased by having previously attached to the apparatus a bulb containing charcoal prepared from coconuts or cherry-stones. If the charcoal is reduced to the temperature of liquid air [− 180° C.], it absorbs nearly all the residual gas and vapours [the Toepler pump will not remove vapours]. Instead of using charcoal, which is likely to explode at low temperatures if its gas content is high, it is better to use dried granular gelatinous silica in the bulb which is cooled, as this substance gives rise to no danger.

In the manufacture of wireless valves and X-ray tubes, mercury vapour pumps are employed to create a very high vacuum in them,

but these pumps can only be used with an auxiliary or 'backing' pump, i.e. the pressure in the apparatus must be low [< 1 cm. of mercury] before they will work. The mercury vapour pump described below is capable of producing an X-ray vacuum when backed by a filter pump, but the best mercury vapour pumps require to be backed by a rotary vacuum pump—cf. the next section. The modern condensation pump was originally designed by LANGMUIR, but nowadays there are many patterns. One designed by WARAN is shown in Fig. 4·21. Mercury is boiled in a vessel A [since the pressure is low, the temperature is seldom above 180° C.] and a mercury vapour jet is formed at C.

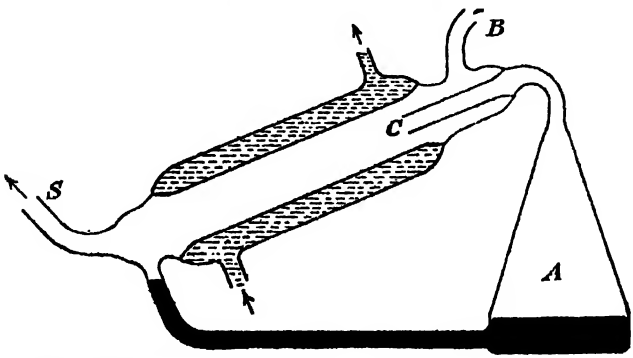


FIG. 4·21.—Mercury Vapour Pump (or Diffusion Pump).

The vessel to be exhausted is connected at B, whilst a water pump is attached to S. Around the wide tube into which the nozzle C projects, there is a water jacket, through which a constant stream of water flows. Consider the state of things in the neighbourhood of the jet. Molecules of mercury vapour and of the gases will tend to intermingle. They are said to diffuse. The mercury vapour, which diffuses towards B, is condensed, whereas the gaseous molecules diffuse towards S and are withdrawn by the water pump. In this way a very low vacuum is reached, but one must not imagine that *all* the molecules have been removed even in the highest vacua which have been produced. There still exist in such vacua about twenty millions of molecules per mm.^{3}.

A Rotary Vacuum Pump.—The pump shown in Fig. 4·22 is designed for the production of a high vacuum and the exhaustion of vessels of large capacity. It works directly from atmospheric pressure and being entirely immersed in oil the leakage of air into the high vacuum is prevented. The pump consists of an outer steel casing, C, through which is bored a cylindrical chamber, D. A shaft, M, runs through this chamber, its axis being parallel to but eccentric from the axis of the chamber. This shaft revolves about its own axis and always touches the periphery of the chamber D at the point E. On each side of this point is a port—one an inlet, F, and the other an outlet, G, which is fitted with a spring-loaded valve, H. In the shaft M is a slot in which two plates, P and Q, are free to slide to and from the axis of the shaft. These two plates are kept apart and their extreme edges forced against the periphery of the chamber D by a series of springs placed at right angles to the axis of the shaft—one of these is shown in sectional view.

The action of the pump is as follows. Let us consider the position shown in the diagram. The shaft M is rotating in an anti-clockwise direction and the effective space between the chamber D and the shaft M is divided into two portions, S and T. As the shaft rotates, remem-

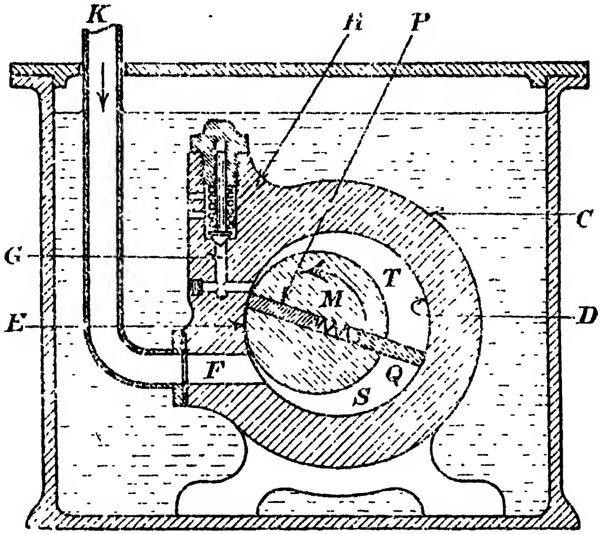


FIG. 4·22.—A Rotary Vacuum Pump.

bering that the plate Q is touching the wall of the chamber, the portion S enlarges and air is drawn in from the vessel to be exhausted through the inlet pipe K. The portion T is getting smaller and any air in it will be compressed. When the pressure is sufficiently great this air escapes through the exhaust valve. Thus the pump will exhaust air from a vessel to which the inlet pipe K is connected.

The Measurement of Low Pressures.—When it is necessary to know the pressure inside a partially exhausted vessel a manometer is used. This consists essentially of a U-tube closed at one end. The closed end is *completely* filled with mercury but there is only a small amount in the other limb of the tube. When the manometer is connected to a vessel from which the gases are being removed gradually, a point is finally reached when the mercury begins to descend in the closed limb of the tube. Finally the difference in level between the mercury surfaces in the two tubes becomes constant and is then a measure of the pressure of the remaining gas in the vessel which is under evacuation. Such manometers possess several disadvantages :—

(*a*) The vacuum in the closed limb is gradually destroyed by gases which creep between the mercury and glass surfaces.

(*b*) If the apparatus suddenly develops a leak the mercury is forced rapidly into the closed limb and the impact is sufficient to cause a fracture of the manometer.

(*c*) The instrument is not sensitive at low pressures.

(*d*) The mercury tends to stick to the glass so that it becomes difficult to observe the true pressure.

The first two disadvantages can be minimized by the use of a device due to WARAN. A small glass reservoir R, Fig. 4·23, is joined by means of capillary tubing to the usual form of manometer. The whole is filled with mercury as before. When the pressure upon the free surface of the mercury is diminished, at some stage the mercury recedes from the point A. If at this stage the instrument is tapped gently, the continuous thread of mercury in the capillary tube is broken and the mercury assumes the position shown in the diagram. The capillary tube space is then an almost perfect void, so that the height BC is a true representation of the pressure at C.

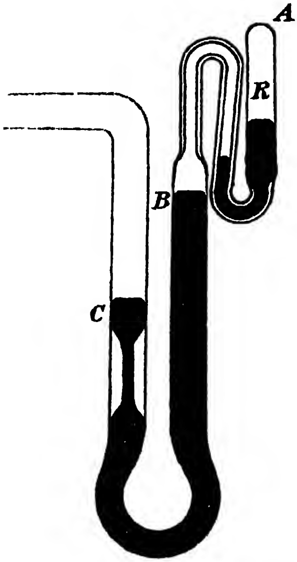


FIG. 4·23.—Manometer with Regenerative Vacuum Device.

After some time gases may make their appearance in the capillary; they are removed by subjecting the manometer to atmospheric pressure, thereby forcing them into R. By constricting the open limb of the U-tube as shown in the diagram, the motion of the mercury is retarded so that a fracture from the causes mentioned above becomes a very remote possibility.

The McLeod Gauge.—Since it is impossible to use a mercury manometer to measure high vacua (such as exist in wireless valves) it is important to discover a means whereby this may be done. McLEOD is responsible for the gauge which is frequently used for this purpose. A bulb A, Fig. 4·24 (*a*), of known volume V, has fixed to its upper extremity a capillary tube DE, the volume of which per unit length is known. The tube BC leads to the apparatus in which it is desired to measure the pressure. A reservoir F contains mercury and is attached to the gauge proper by means of pressure tubing G. When the reservoir F is lowered through a distance greater than that equal to the barometric height (say 80 cm.) below the level B, then A is in direct contact with the exhausted vessel, and is therefore filled with gas at a pressure p, which is the pressure to be determined. When F is raised, the mercury divides at B and entraps a volume V of gas at pressure p; by raising F still more this gas can be compressed into the capillary DE. To derive a value for p the mercury in C may be adjusted until it is level with the closed end D of the capillary tube. Then the pressure of the gas in DE is measured by h, where h = DE. Now Boyle's law [cf. p. 85] states that the product of the pressure (p) and the volume (V) is constant for any given mass of gas at constant temperature. Applying this to the mass of gas entrapped in the capillary, we have

$$pV = hx,$$

where x is the volume corresponding to the length DE of the capillary tube. Whence

$$p = \frac{hx}{V} = \frac{h^2v}{V},$$

if v is the volume per unit length of the capillary.

If such a gauge is to be reliable the enclosed gas must be dry, for water vapour does not behave like an ideal gas.

In the more recent forms of this instrument a piece of glass tubing of the same diameter as that used for DE is sealed in parallel with the side tube C as shown in Fig. 4·24 (*b*). When reading the difference in levels of the mercury in the tube E and that leading to the vacuum, it is the levels in E and this other tube which must be recorded. This is because the surface tension of mercury is such that it is depressed in narrow tubes to an extent depending on the diameter of the tube. The effect is eliminated, however, by using tubes of the same diameter.

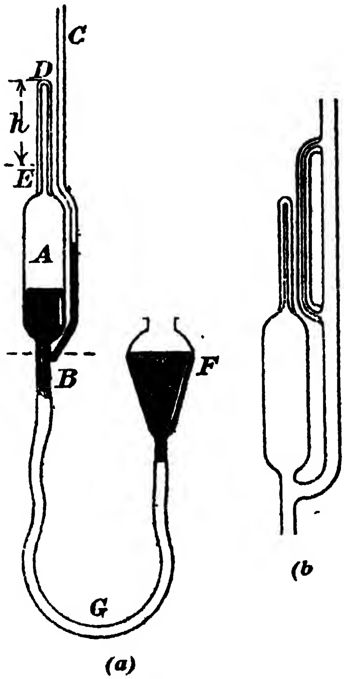


Fig. 4·24.—McLeod Gauge for Measuring Low Pressures.

The Absorption of Gases.—The process of obtaining a high vacuum is by no means as simple as the above remarks would indicate. It is found that after a certain time, depending on the pump and the nature and size of the vessel to be exhausted, the pressure ceases to be reduced. This is because gases are evolved from the surfaces of all substances when the external pressure is very low. The rate at which these gases is expelled is greatly increased when the temperature of the surface is raised. The vessels to be exhausted are therefore heated cautiously with a gas flame and the pumping continued.

If, as in a wireless valve, there is some metal to be degassed, it is subjected to a heavy electron bombardment. We shall learn later that electrons are emitted when a metal is heated to high temperatures. A filament is therefore placed near the metal (or the filament of the valve used) and its temperature raised electrically. A large positive potential is then applied to the metal, while the filament is earthed at one point. The electrons are attracted to the metal and strike it with considerable velocity. They lose their kinetic energy which appears as thermal energy [heat], and it is this energy which is responsible for the liberation of the occluded gases in the metal.

EXAMPLES IV

1.—Calculate the mass of lead, density 11·3 gm. cm.$^{-3}$, which must be attached to 105 cm.^{3} of wax (density 0·86 gm. cm.$^{-3}$) in order that the apparent mass may be zero when the whole is placed in a liquid whose density is 1·04 gm. cm.$^{-3}$.

2.—A U-tube contains mercury, density 13·6 gm. cm.$^{-3}$. A liquid whose density is 1·23 gm. cm.$^{-3}$ is poured into one limb so that the difference between the mercury levels is now 3·67 cm. What is the length of the column of liquid? Can you make any statement concerning the mass of the liquid which has been added?

3.—The height of a water barometer is 34 ft. Find the pressure in atmospheres 1 mile below the surface of sea water (density 1·026 gm. cm.$^{-3}$). Also express this pressure in ton.-wt. ft.$^{-2}$ [1 cu. ft. of water has a mass of 1000 oz.].

4.—Find the pressure due to a column of air 1 mile high if the density of the air is uniform and equal to 0·00129 gm. cm.$^{-3}$. Describe how a barometer may be used to determine the height of a mountain.

5.—A rectangular tank measures 4 ft. by 3 ft. at the base. It is filled with water to a depth of 8 in. What is the depth when a stone (1 ft. cube) is dropped into the tank?

6.—What do you understand by the principle of flotation? An iron cylinder 12·0 in. long floats vertically in mercury. The densities of iron and mercury are 7·8 and 13·6 gm. cm.$^{-3}$ respectively. Calculate the length of iron immersed.

7.—Define the term density. How would you proceed to determine the density of a powder such as plaster of Paris?

8.—How would you determine the density of a newly-laid egg?

9.—Sketch and describe the experimental arrangement you would use in order to obtain a good vacuum. How would you measure the final pressure obtained?

10.—A piece of glass tubing sealed at both ends has a mass 18·26 gm. If the density of glass is 2·63 gm. cm.$^{-3}$, calculate the volume of the air space enclosed in the bulb if the whole has an apparent mass of 6·37 gm. in water.

11.—The space above a mercury column contains some air. The mercury column is 28·40 in. long and the space above is 3·05 in. long. This tube is then pushed downwards into mercury so that the column is 28·14 in. whilst the air space is 2·34 in. What is the true height of the barometer?

12.—What mass of lead, density 11·3 gm. cm.$^{-3}$ must be added to a block of Balsa wood 3·26 cm. × 8·40 cm. × 9·62 cm., and density 6·0 lb. per cu. ft., so that it will just float in water? [1 lb. = 453·6 gm., 1 ft. = 30·48 cm.]

13.—A pellet of mercury, density 13·59 gm. cm.$^{-3}$ mass 5·278 gm., has a length 20·4 cm. when introduced into a narrow tube. What is the average radius of this tube? Some liquid is then placed inside the tube and the length of the column is 18·9 cm. What is the density of the liquid if its mass is 0·467 gm.?

14.—What is meant by the statement that the pressure of a coal gas supply is 12 cm. of water? If the pressure of the gas supply at ground-level is 12 cm. of water what will be the pressure of the supply at the top of a building 25 metres high if the relative densities of gas, air, and water are as 1 : 2 : 1,450?

15.—Explain the conditions on which floating depends. A cork of

specific gravity 0·25 floats in sea-water of specific gravity 1·25 with 10 cm.^{3} above the surface. Calculate the total volume of the cork.

16.—Define density. If the density of glass is 2·265 gm. cm.$^{-3}$, express its density in terms of the lb. and yard when these are the units of mass and of length respectively. [1 lb. = 453·6 gm., 1 in. = 2·540 cm.]

17.—If you were supplied with some turpentine and some ice, describe how you would determine the density of the ice without using any form of balance or 'weights.'

18.—A body 'weighs' 86·0 gm. in air, 72·4 gm. in one liquid and 63·9 gm. in another liquid. In a mixture of these liquids it 'weighs' 67·1 gm. Calculate the proportion in which the liquids have been mixed.

19.—A solid whose density is 12·4 gm. cm.$^{-3}$ is weighed in air. It is found that its mass is 284 gm. when brass weights having a density 7·8 gm. cm.$^{-3}$ are used. If the density of air is 1·25 gm. litre^{-1}, calculate the error due to neglecting the buoyancy of the air.

20.—A cylinder of 0·3 cm.^{2} cross-section is loaded at one end and the whole has a mass of 6·43 gm. In water it is found that 1·8 cm. project above the surface. Calculate the amount of this projection when the cylinder floats upright in a liquid whose density is 1·37 gm. cm.$^{-3}$.

21.—Describe a modern form of barometer. What is a bar? Calculate the number of bars in one standard atmosphere.

22.—The pressure at a depth of 100 ft. in a fresh-water lake is three times the pressure at a depth of 11 ft. Determine the height of the mercury barometer in cm. [Density of mercury = 13·6 gm. cm.$^{-3}$.]

23.—A column of mercury is placed at the middle of a uniform glass tube and both ends of the tube are closed when the tube is horizontal, and the pressure everywhere 76 cm. of mercury. The tube is then placed vertically and it is found that the length of the tube occupied by the air above the mercury is twice as great as that occupied by air below the mercury. What is the length of the mercury column?

24.—If a series of observations of the volume, V, of dry gas enclosed in a Boyle's law apparatus and the excess pressure (p) inside the apparatus were made, explain how the atmospheric pressure may be deduced from a graph showing the relation between p and $\frac{1}{V}$.

25.—Describe how you would proceed to verify Boyle's law. The height of a faulty barometer which has a little air in the space at the top of the mercury column is 28·6 in. when the barometric height is 29·1 in., and 29·2 in. when the true height is 30·1 in. Calculate the barometric pressure when the instrument indicates 28·9 in.

26.—The following figures are taken from the treatise in which Boyle published an account of one of his experiments made to determine the relation between the pressure and volume of *a given parcel of air* at room temperature. Use them to find a value for the height of the barometer on the day when this experiment was made.

Length of tube occupied by air (inches)	$11\frac{1}{2}$	$10\frac{1}{2}$	9	8	$5\frac{3}{4}$	$4\frac{1}{2}$	3
Excess pressure of the air inside the tube over atmospheric pressure outside (inches of mercury)	$1\frac{7}{16}$	$4\frac{6}{16}$	$10\frac{2}{16}$	$15\frac{1}{16}$	$32\frac{3}{16}$	$48\frac{12}{16}$	$88\frac{7}{16}$

CHAPTER V

CONCERNING THE NATURE OF FLUIDS

The Brownian Movement.—To an observer standing on the landward side of a breakwater the nature of the tempestuous seas beyond that breakwater can be inferred from the rolling and pitching motions of the ships which will be more excessive than usual. To the eye, aided by the most powerful of microscopes, the motion of molecules cannot be made visible. If, however, some small particles of gamboge suspended in a liquid are observed with the aid of a microscope, it will be found that these particles are always moving, not in any fixed direction, but in all random directions. The actual motion of a particular particle is very irregular, and perhaps the most striking feature of this phenomenon is that the motion never ceases. This phenomenon, discovered by an English botanist Brown early in the last century, has been observed in liquids contained in the enclosed cavities of some varieties of quartz, and these cavities and the liquids in them will have been there for thousands of years. It has been concluded that this eternal motion of the suspended particles cannot be due to any external agencies, but must be attributed to the movements of the molecules which constitute the liquid.

The Brownian motion can also be detected in ***collosol oil of iodine.*** This substance is applied to the patient's skin in cases where it is necessary to alleviate the pain due to rheumatism, sciatica, etc. The small particles of iodine are participating in this so-called Brownian movement, and consequently they are able to pass very readily through the skin and into the body.

Diffusion.—Let a quantity (say 25 $cm.^3$) of a concentrated nickel (or copper) sulphate solution be placed at the bottom of a tall glass cylinder, the remainder of the vessel being filled with water. A glass cover prevents evaporation. Such a coloured substance is chosen so that the movements of the resulting solution may be observed easily. At first the line of demarcation between the water and the solution is well defined, but it becomes obliterated after a lapse of several days. The dissolved substance has moved upwards against the pull due to gravity, i.e. it has moved to a region where

the concentration of the salt in solution was less. The rate at which this transference of the dissolved substance takes place is very slow. It would be very difficult to explain this phenomenon if the molecules of the liquids were not in a state of continual irregular motion. The molecules of the dissolved substance—or, in the case of electrolytes, the ions in the solution—behave, in this respect, like the molecules of a gas, and the process by which molecules in different solutions move from regions of higher to those of lower concentration, or the molecules of one gas intermingle with those of another is called ***diffusion***. In a gas the molecules are at relatively large distances from one another and so are free to move. The molecules of the dissolved body in a solution may be regarded, for some purposes, as being distributed throughout the solvent ; the solvent has merely made it possible for the constituent molecules of the dissolved body to occupy a space much beyond the original confines of the crystal.

The Diffusion of Salts in Aqueous Solution.—In 1850 GRAHAM published his first paper on the diffusion of salts in solution, and in 1882 a further study was made by SCHEFFER. In principle the apparatus they used is shown in Fig. 5·1 (*a*). A small glass cylinder, A, rests on two horizontal glass rods supported inside a larger glass vessel, B. A is nearly filled with the solution under investigation, and a cork, C, floats centrally on the liquid. A vertical knitting needle attached to this cork can move upwards

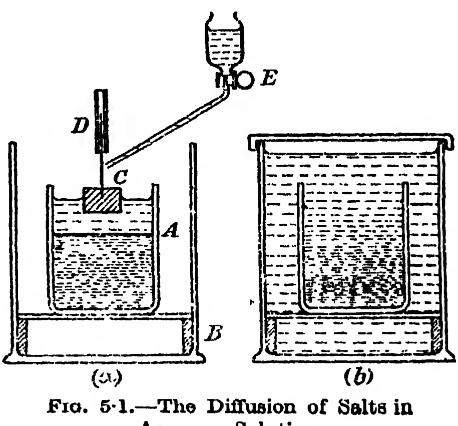


FIG. 5·1.—The Diffusion of Salts in Aqueous Solution.

in a narrow glass tube, D, held in position by a clamp and stand (not shown). By this means the cork is kept in a central position. Water is contained in the dropping funnel, E, and it is allowed to drop on to the top of the cork, which has been thoroughly wetted, at the rate of about three drops per second. A layer of water soon appears on top of the solution, and when the cork is clear of the solution, it may be removed, and the vessel, A, completely filled with water. The whole of A is then surrounded by water as in Fig. 5·1 (*b*). The temperature is kept constant to avoid convection currents. At first there is a distinct boundary between the solution and the water. As a result of the process called

diffusion this well-defined boundary soon disappears. By determining the amount of solute which had escaped from the inner vessel into the outer one, it was found

(i) the rate of diffusion depends on the nature of the dissolved substance, so that the ratio of the amounts of two substances present in a solution may alter on account of diffusion,

(ii) the rate of diffusion is directly proportional to the concentration of the dissolved substance,

(iii) a rise in temperature augments the rate at which diffusion takes place.

Fick's Law.—Four years after the publication of Graham's first paper on diffusion, FICK, guided by Fourier's work on the conduction of heat, enunciated the following law. ***The mass, m, of a substance in solution passing across an area A per second is directly proportional to the rate at which the concentration, c, of the dissolved substance diminishes in a direction at right angles to the plane of the area A.*** In symbols

$$\frac{m}{A} = -D\frac{dc}{dx},$$

where D is the coefficient of diffusion, and $\frac{dc}{dx}$ is the rate at which the concentration increases with the distance x.

The Passage of Gases through Porous Bodies.—The diffusion of two gases is not prevented but only hindered when a thin porous wall or membrane separates them, but the actual rate at which the gases intermingle depends upon several factors. If the pores through which the gas passes are short in comparison with their diameters the gas flow is similar to that of water through a hole in the side of a thin-walled container. This process is known as ***effusion***. The velocity of effusion is proportional to $\sqrt{\frac{p}{\rho}}$, where p and ρ are the excess pressure of the gas above that of the surrounding air and the density of the gas or gas mixture passing through respectively. When the pores are reduced in diameter the flow of gas, for a given difference in pressure between the ends of the tube or pore and provided that the pressure difference is not so large that turbulent motion ensues, is controlled by the viscosity of the gas [cf. p. 130]. In both these instances the gas passes through as a whole so that if it were a mixture of gases no partial separation would be effected. Conditions are very different, however, when the pores are so fine that their diameters are comparable with those of the gas molecules. GRAHAM, who first investigated these phenomena about 1840, discovered that the rate of diffusion at a given temperature was directly proportional to the difference

in pressure between the two sides of the membrane, and inversely proportional to the square root of the density of the gas. This is known as ***Graham's Law of diffusion for gases.***

Hence, for a given pressure difference, hydrogen diffuses four times as quickly as oxygen through the same membrane, since, under the same conditions, the density of a gas is directly proportional to its molecular weight. This implies that if an oxygen-hydrogen mixture is introduced under pressure into a porous vessel the mixture passing through will be four times as rich in hydrogen as in oxygen.

The diffusion of gases through porous media may be investigated

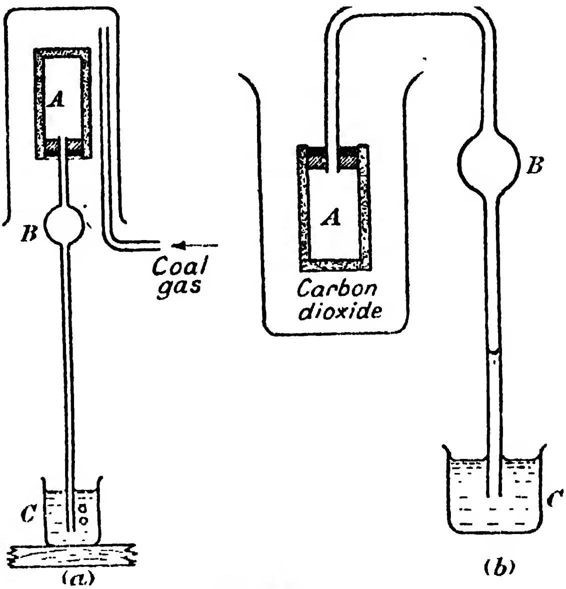


FIG. 5·2.

experimentally with the aid of the apparatus shown in Fig. 5·2 (*a*). A glass tube 60 cm. long and 0·5 cm. wide passes through a cork from a porous pot A to a vessel containing coloured water. The cork is pushed well within the pot and covered with sealing-wax to make the joint air-tight. A large jar is held over the pot and coal gas introduced into it. Bubbles of gas at once escape from the tube at C showing that the pressure in A is tending to increase. If the jar is removed the stream of bubbles at once ceases and the liquid rises in the tube. The bulb B is sufficiently large to prevent liquid from reaching A. In the first part of this experiment the coal-gas passes more rapidly into the pot than the air inside can escape, so that the pressure rises. In the second part, the coal-gas

which has found its way into the pot diffuses out more rapidly through the walls than the air does inwards so that the pressure inside is reduced.

A similar experiment may be made using carbon dioxide instead of coal gas. For this purpose the apparatus is arranged as in Fig. 5·2 (*b*). A jar containing the carbon dioxide is placed so that there is an atmosphere of the gas round the porous pot. The liquid rises in the tube, showing that air is diffusing more rapidly from the pot than carbon dioxide is diffusing inwards. When the jar is removed, the pressure inside the apparatus increases and, depending on the relative amount of carbon dioxide which has entered the pot, a bubble of gas may escape from the tube immersed in the liquid.

The Diffusion of Solids.—Diffusion in solids has been investigated by Sir Roberts-Austen, who placed an alloy of lead and gold (5 per cent. gold) in contact with a piece of lead, the two surfaces in contact being accurately plane and held together under pressure. The whole was heated at 165° C. for one month. On analysing various sections it was found that diffusion had taken place. The experiments were repeated at room temperature when it was observed that diffusion still occurred, only at a diminished rate.

The diffusion of one solid into another finds an important application in the 'cementation' process of converting iron into steel. The iron is placed in intimate contact with powdered carbon and then heated. The depth to which carburization takes place depends upon the temperature and time of heating.

Osmosis.—When red blood corpuscles are placed in water they expand rapidly and ultimately burst, but if they are placed in a strong salt solution they shrivel up. This phenomenon is characteristic of the membranes surrounding many animal and vegetable cells, for these allow water to pass through freely but retard or entirely prevent the passage of solids. ***Osmosis*** is the name given to this spontaneous passage of a liquid through a membrane. Its effects were first observed by the Abbé Nollet in 1748, but it was left to a botanist, Pfeffer, to investigate it quantitatively. A piece of wet parchment paper is stretched over the end of a large thistle funnel and when nearly dry it is coated with glue along the boundary. The inverted funnel is partly filled with a solution of sodium chloride, cane-sugar, or some other substance, and immersed in water [cf. Fig. 5·3]. After standing for some time the level of the solution will have risen considerably; water must have passed through the parchment into the solution. This statement is not complete, for water will have passed from the solution into the water in the beaker at the same time as water passed from the beaker into the solution. This osmotic flow arises from the bombardment of the molecules upon the membrane; on the one side there are only molecules of water arriving at the membrane, whilst on the

other hand there are molecules of water and solute as well. Now such membranes are only slightly permeable to dissolved salts and the resultant effect is that more water molecules pass in one direction than the other.

An osmotic flow of the solvent is also observed when a membrane separates two solutions of the same nature but differing in concentration. The flow of solvent is such that the concentrations of the solutions tend to become equal, i.e. there is an excess of solvent passing from the weaker to the stronger solution.

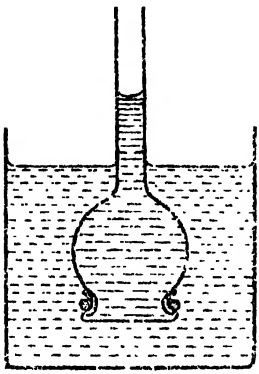


FIG. 5·3.—Osmosis.

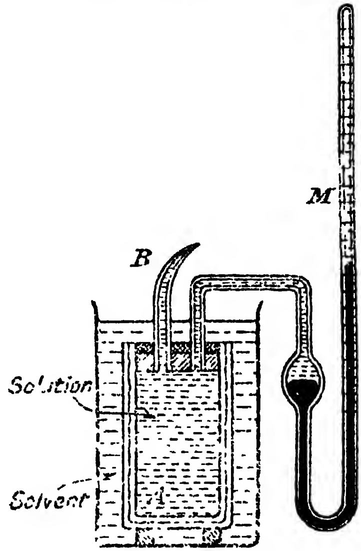


FIG. 5·4.—The Measurement of Osmotic Pressure.

Semipermeable Membranes.—A membrane which permits the solvent but not the solute to pass through it is termed a semipermeable membrane. One of the best-known membranes of this class is copper ferrocyanide.

Experiment. Place a weak solution of potassium ferrocyanide in the bottom of a beaker and when it has ceased to move introduce a strong solution of copper sulphate so that it lies below the ferrocyanide solution. A thin gelatinous precipitate of copper ferrocyanide is formed: it separates the two solutions. The membrane does not increase in thickness since the dissolved substances cannot pass through it, but after the lapse of about two hours it will be seen that the membrane has a distinct bulge upwards. This proves that more water passes downwards than flows upwards, and hence that the copper solution has the greater osmotic pressure.

Osmotic Pressure.—The membrane of copper ferrocyanide prepared in the above experiment is too fragile to support more than a

small pressure difference, but its strength is very considerably increased if it is produced in the walls of a porous pot. To prepare this membrane the porous pot is boiled in distilled water for several hours to remove air bubbles. A 0·25 per cent. solution of copper sulphate is then placed inside the pot and a 0·21 per cent. solution of potassium ferrocyanide outside. Each solution should reach very nearly to the top of the pot. Diffusion occurs and the two dissolved substances meet inside the walls of the pot where a membrane of copper ferrocyanide is formed. This process should be allowed to continue without interruption for two days. The pot thus prepared is boiled in several changes of distilled water and is then ready for use. If allowed to become dry it should be boiled for several hours to expel all air again.

If such a pot, provided with a rubber bung carefully waxed in position and provided with a long capillary tube, is filled with a saturated solution of cane sugar and then immersed in water, the change in level of the liquid in the capillary is very rapid. After several days a tube 1 mm. in diameter must be several metres long if the liquid is not to exude from it. This spontaneous differential flow of liquid through the membrane can be completely stopped by the application of a suitable pressure; the flow is reversed if the pressure is increased beyond this value.

Definition.—***That pressure which must be applied to a solution to prevent the spontaneous differential flow of liquid through a semipermeable membrane separating the solution and solvent is termed the osmotic pressure of the solution.***

To determine the osmotic pressure of a weak aqueous solution the apparatus shown schematically in Fig. 5·4 may be used. A mercury manometer, M, with one limb closed and containing air, or better, nitrogen, is connected to the porous pot, A, containing the solution. This solution is introduced through the tube B, which is afterwards hermetically sealed and the air in the connecting tubes displaced by some of the solution so that temperature changes do not affect the volume between the pot and the gauge. Water enters the solution and the pressure inside the pot increases. Ultimately this pressure ceases to change and this constant pressure is the osmotic pressure of the solution. It is calculated from the change in volume of the gas (air) in the closed limb of the manometer. The serious objection to this method lies in the fact that the water entering the solution changes the concentration of the latter so that the readings do not correspond to the osmotic pressure of the original solution: neither do they to the final solution, for its concentration is not uniform and it is the concentration of the solution in the immediate vicinity of the mem-

brane which determines the osmotic pressure which is measured. It is better to measure the external pressure which must be applied to the solution to prevent the passage of the solvent. Such methods must always be used for concentrated solutions. LORD BERKELEY and HARTLEY have developed this method, but their apparatus is too complicated for a detailed description here.

The Fundamental Laws of Osmotic Pressure.—(*a*) At constant temperature the osmotic pressure of a dilute solution is directly proportional to the concentration of the solute in the solvent, i.e. it is inversely proportional to the volume of the solvent containing a given mass of dissolved substance.

(*b*) The osmotic pressure of a dilute solution is directly proportional to its absolute temperature.

The analogy between these two laws and those of Boyle and of Charles is very apparent: in fact, the osmotic pressure of a dilute solution is the pressure which the dissolved substance would exert if it existed as an ideal gas occupying the same volume and being at the same temperature as the solution.

The above laws apply to dilute solutions of non-electrolytes, but experiment shows that solutions of electrolytes have higher osmotic pressures than they would indicate. This is explained by the fact that such substances exist as ions when they are in solution.

The laws of osmotic pressure may be symbolized by the formula

$$pv = k\mathrm{T},$$

where p is the osmotic pressure, v the volume of solution containing 1 gm. of the solute, T is the absolute temperature, and k is a constant. VAN'T HOFF showed that the constant k in the above 'osmotic equation' had the same value as the constant $\mathscr{R}$ in the characteristic equation for an ideal gas.

If one mole of a substance of molecular weight M is dissolved in a volume V cm.^{3}, then $\mathrm{V} = \mathrm{M}v$, so that the characteristic equation becomes

$$p\mathrm{V} = \mathrm{M}\mathscr{R}\mathrm{T}$$

Now it is found that $\mathrm{M}\mathscr{R}$ is a constant for all substances: it is denoted by R and is known as the universal gas constant. Thus

$$p\mathrm{V} = \mathrm{RT}$$

If Ω is the volume of a solution containing N moles of dissolved substance, $\mathrm{NV} = \Omega$, so that

$$p\left(\frac{\Omega}{\mathrm{N}}\right) = \mathrm{RT}.$$

If $C = \frac{N}{\Omega}$, the concentration in mole. cm.$^{-3}$, then

$$\frac{p}{C} = RT.$$

If c is the concentration in gm. cm.$^{-3}$, $c = MC$, so that

$$\frac{p}{c} = \frac{RT}{M}.$$

Thus, if the osmotic pressure, in absolute units, of a solution at temperature T and concentration c is known, it is possible to determine the molecular weight of the dissolved substance.

Osmotic Pressure and the determination of Molecular Weight.—It is customary in experimental work on osmosis to measure the pressure in atmospheres and to consider the volume in cm.^{3} occupied in solution by 1 gram-molecule (or 1 mole) of the dissolved substance. The characteristic equation for an ideal gas, when the pressure is measured in atmospheres and 1 gram-molecule occupying a volume V is considered, then becomes

$$PV = \bar{R}T,$$

and $\bar{R}$ is a universal constant for all gases. It must be noted, however, that $\bar{R}$ is different from the universal gas constant R which appears in the ideal gas equation $pV = RT$, where the pressure p is expressed in absolute units. It is known that 1 gram-molecule of a gas at S.T.P. occupies 22,415 cm.^{3}. Hence

$$1 \times 22{,}415 = \bar{R} \times 273,$$

$$\text{or } \bar{R} = 82{\cdot}06 \text{ cm.}^3 \text{ atmos. deg.}^{-1} \text{ K. mole.}^{-1}$$

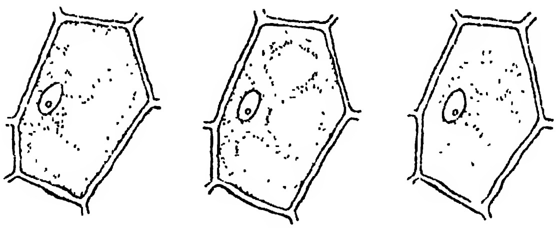
This enables us to calculate a value for the osmotic pressure of a non-electrolyte in solution or, knowing the osmotic pressure, to determine a value for the molecular weight of the dissolved substance. Let m gm. of a substance of molecular weight M be dissolved in 100 cm.^{3} of water at $\theta°$ C. Then the number of gram-molecules in this volume is $\frac{m}{M}$, so that 1 gram-molecule would occupy $\left(\frac{M}{m} \times 100\right)$ cm.^{3}. Let P be the required pressure in atmospheres. Then

$$P \times \left(\frac{M}{m} \times 100\right) = \bar{R} \times (273 + \theta)$$

$$\therefore\ P = \frac{0{\cdot}821 \times (273 + \theta) . m}{M} \text{ atmosphere.}$$

$$= \frac{0{\cdot}821 m(273 + \theta)}{M} \text{ atmosphere.}$$

Isotonic Solutions and Plasmolysis.—If the variation with concentration of the osmotic pressure of one aqueous solution at a constant temperature has been investigated experimentally, the unknown osmotic pressure of another aqueous solution may be found in the following way. The strength of the first (or standard) solution must be adjusted until its osmotic pressure equals that of the unknown solution; the two solutions are then said to be ***isotonic*** with each other. To carry out such an experiment a convenient semi-permeable membrane must be available. DE VRIES, a Dutch botanist, in 1888 used the cells of the leaves of certain plants, among which he mentions those of *Tradescantia discolor, Begonia manicata,* and *Curcuma rubricaulis.* Such cells consist of a mass of protoplasm (living matter) containing sap vacuoles separated from the protoplasm by the so-called ***inner plasma membrane,*** and surrounded by a cellulose wall which is sufficiently



(a) (b) (c)

// = Cellulose wall · = Protoplasm with sap vacuoles. O = Nucleus.

FIG. 5·5.

strong to withstand forces tending to change its shape, i.e. its changes in form are only minute. This wall is separated from the general mass of protoplasm by the ***outer plasma membrane.*** The whole of the contents of such a cell is termed the ***protoplast.***

Now the vacuoles contain the cell sap in which dissolved substances exist. It is not known for certain whether or not the plasma membranes are the controlling semi-permeable membranes through which water passes to the vacuoles, or whether the whole lining of protoplasm acts in this way. The latter assumption is generally adopted as it simplifies the discussion.

If, therefore, these cells are immersed in a solution having an osmotic pressure equal to their own, the cell, viewed under a microscope, will present its normal appearance [Fig. 5·5 (*a*)]. If the cell is placed in a solution having a greater osmotic pressure than its own, water will pass from the cells into the solution; the protoplast will shrink and the cell will appear as in Fig. 5·5 (*b*), or finally as in Fig. 5·5 (*c*). If the cell is placed in a solution the osmotic

pressure of which is less than its own then water will pass into the cell, but this will only be very slightly distended on account of the relatively strong cellulose wall which forms the external boundary of the cell.

In order to find a solution which shall have an osmotic pressure equal to that of the cell, experiments are first made with a solution having an osmotic pressure greater than that of the cell. The solution is then diluted gradually until the protoplast just maintains its normal form. When this occurs the solution in the cell and the one in which the cell is immersed, each exert the same osmotic pressure, i.e. they are ***isotonic*** with one another. The above method of determining osmotic pressure either of the solution in the vacuole of a leaf, or of an unknown aqueous solution, is referred to as the ***plasmolytic*** method.

Dialysis.—In his famous researches on the phenomenon of diffusion, GRAHAM found that some substances (mineral acids and salts), the so-called crystalloids, were able to pass through certain semi-permeable membranes. The other type of substance (gum, for example) is known by the name of colloid. The line of demarcation between the two types is not sharp, some substances behaving like crystalloids or colloids according to the nature of the solvent in which they are dissolved. The classical example is that of sodium stearate, $C_{17}H_{35}.COONa$, which acts as a colloid when an aqueous solution is made, whereas it exhibits the properties of a crystalloid when in alcoholic solution. Crystalloids are such that when they are dissolved in water, they produce a diminution of its saturation vapour pressure, a fact which is revealed by the lowering of the freezing-point and the raising of the boiling-point of the water; on the other hand, colloids produce no appreciable effect. Whenever a colloid is made it almost invariably contains a quantity of the crystalloid from which it has been prepared. The separation of these substances is carried out by means of a process known as 'dialysis.' The mixture is placed in a cylindrical vessel, the bottom of which consists of parchment paper. The whole is placed in a liquid medium capable of dissolving the crystalloid. The crystalloid diffuses through the membrane until the concentration of this substance is the same on both sides of the medium. Frequent renewals of the solvent are therefore made, and in this way a colloid, free from crystalloids, is obtained.

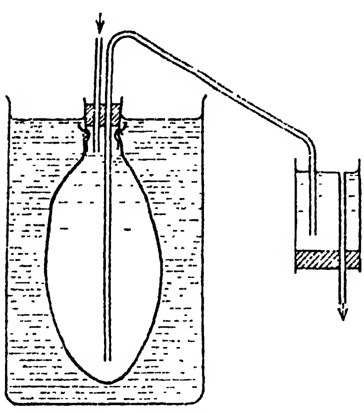


FIG. 5·6.—Apparatus for use in 'Hot Dialysis.'

The membranes which are used for dialysis are gold-beaters' skin, fish bladder, and parchment paper. The speed at which dialysis takes

place rapidly increases with rise in temperature, and in order to effect this increase the hot dialyser shown in Fig. 5·6 may be employed. The colloid from which the crystalloid is to be removed is placed in a two-litre beaker. This is heated with the aid of a suitable burner. A membrane is attached by string to a cork suitably bored and fitted with two glass tubes to allow distilled water to pass into the bag which the membrane forms, any excess being removed by means of the automatic device indicated. This excess water carries with it the crystalloid which has passed through the membrane.

Surface Energy and Surface Tension.—Everyone will have noticed that when a small amount of liquid is brought into close contact with a solid, the liquid either spreads itself over the surface of the solid, or else collects itself into small drops, and that most liquids tend to rise in capillary tubes to a distance above the surface of the liquid in the containing vessel, whereas some, such as mercury and molten metals, act in an exactly opposite way. To explain these phenomena it has sometimes been maintained that the surface of a liquid must be endowed with some peculiar property, e.g. the surface may be skin-like. LANGMUIR and N. K. ADAM have shown that all these properties of liquids can be attributed to molecular happenings inside the liquid. The hypothesis that the surface of a liquid has a skin-like structure has been superseded by these more modern views. The fact that the molecules of a liquid are free to move has been confirmed by experiments on Brownian movement. These molecules must be very closely packed together, for experiment has shown that a liquid resists forces tending to compress it, even if the forces are enormously large. Since the molecules are so close together, the forces of attraction between neighbouring molecules in liquids must be very large. When, however, a molecule is at the surface of the liquid it will not be attracted equally in all directions, for there is no liquid above it. In consequence of this such molecules will tend to move towards the interior of the liquid. Since the molecules occupy space, i.e. there is a definite number per unit volume, the surface tends to diminish in area. In support of these remarks we have the fact that liquids tend to assume that shape which has the minimum area for a given volume. If a drop is subject to other forces comparable with those discussed above its shape will be slightly distorted from that having the minimum area, e.g. a rain-drop hanging from a window-pane.

In virtue of these forces, directed inwards, molecules at the surface will possess a certain amount of energy due to position. The amount of this energy per unit area is termed the ***surface energy***. The surfaces of both liquids and solids possess surface energy but it is only when the surface is mobile that its effects become apparent. The fact that a liquid surface is the seat of

potential energy manifests itself very vividly when a soap film is ruptured (with the aid of a pointed piece of filter paper, for example), for the liquid is projected in all directions with a considerable velocity, i.e. the potential energy has been converted into kinetic energy.

Let a liquid film be formed between two limbs of a bent wire, BAC, Fig. 5·7, and a horizontal straight wire, XY, placed across them. Suppose that a force, F, acting normally to XY is necessary to maintain equilibrium when the film is vertical. Then F must be balanced by a force on the wire due to the film. Suppose γ is the magnitude of this force per unit length of the wire. If the length of XY is l, the total force on the wire from the above cause is $2\gamma l$, the factor 2 being introduced since the film has two sides. Hence

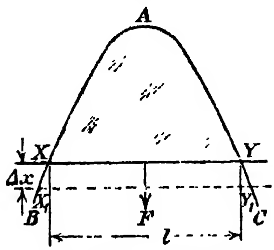


FIG. 5·7.—Surface Tension.

$$F = 2\gamma l.$$

γ is termed the ***surface tension*** of the liquid.

[It should be noted that if parallel wires are used for the purpose of forming a film between them, the system is unstable. For example, if F, is too large, the force $2\gamma l$ never becomes sufficient to balance F for l remains constant. The instability does not matter as far as theory is concerned, but with the stable arrangement here adopted a rough estimate of γ may be made. If the weight of the wire is not sufficient it may be loaded. Then $F = mg$, where m is the total mass of the wire and its load.]

On the Relation between Surface Energy and Surface Tension.—Again consider Fig. 5·7. Let XY move through a small distance δx to a parallel position X_1Y_1 ,the external force on the wire being F. Now when a film is stretched in this way its temperature falls unless heat (thermal energy) is communicated to it. We shall suppose that the heat necessary to restore the film to its original temperature has been supplied.

If ε is the surface energy of the film, i.e. the potential energy per unit area of the surface, the increase in potential energy of the 'surface' molecules is $(2l.\delta x)\varepsilon$, the factor 2 being introduced since the film has two surfaces. The work done by the stretching force is $F.\delta x$. These two quantities cannot be equated, however, for heat has been communicated to it from external bodies. If δQ is the heat (thermal energy) supplied to restore the temperature of the film to its original value, we have

$$(2l.\delta x)\varepsilon = F.\delta x + \delta Q.$$

Now the force F is equal and opposite to the pull of the film on the wire XY, when the film is in equilibrium. If γ is the pull per unit length, then $2l.\gamma = F$, the above equation becomes

$$2l.\delta x.\varepsilon = 2\gamma l.\delta x + \delta Q,$$

or

$$\varepsilon = \gamma + \left(\frac{\delta Q}{2l.\delta x}\right).$$

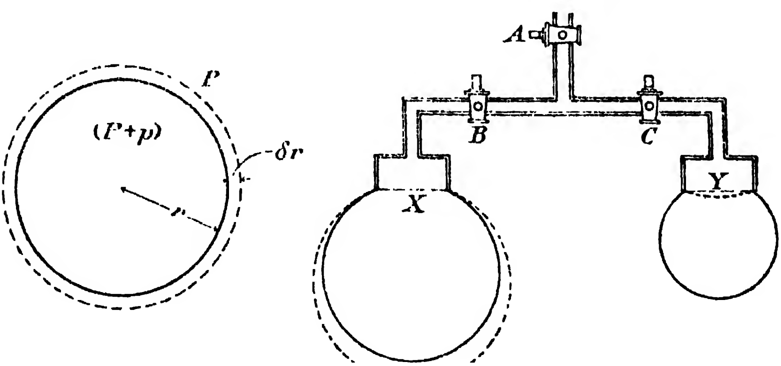
This may be written

$$\varepsilon = \gamma + \eta,$$

where $\eta = \dfrac{\delta Q}{2l.\delta x}$, the thermal energy supplied per unit increase in area of the film.

Now the force γ exerted on each unit length of the wire is called the ***surface tension*** of the liquid and the above shows that the intrinsic surface energy of a liquid is really the sum of two quantities—a 'thermal' part denoted by η, and a 'mechanical' part γ, or $\varepsilon - \eta$; we see, therefore, that the surface tension is equal to the 'mechanical' part of the surface energy. HELMHOLTZ called this 'mechanical' part of the surface energy the free energy of the surface, or the surface free energy. It will be noted also that the increase in the total free surface energy of a surface is equal to the external work done on that surface, providing heat is supplied to keep its temperature constant.

The Pressure Difference across a Spherical Surface.—Let r, Fig. 5·8 (a), be the radius of a spherical bubble of gas in a liquid. Let P be the pressure outside the bubble. We have to show that



(a) FIG. 5·8. (b)

the pressure inside is equal to P + p, where p is a quantity to be determined. For this purpose let r become $r + \delta r$, where δr is a very small quantity, in fact so small that the pressure inside is

not altered thereby. Moreover, let heat be supplied to the film so that its original temperature is restored. The area of the curved surface has increased from $4\pi r^2$ to $4\pi(r+\delta r)^2$. If γ is the surface tension of the liquid or, as we have just seen, its free surface energy per unit area, the increase in free surface energy is $4\pi\gamma[(r+\delta r)^2 - r^2] = 8\pi\gamma r \,.\, \delta r$, since $(\delta r)^2$ may be neglected. This is equal to the work done in expanding the bubble. Since pressure is defined as the force per unit area, the total force acting on the inner surface of the bubble is $4\pi r^2(P + p)$, while that on the outer surface is $4\pi r^2 P$. Since these forces are opposed to one another the net work done on the film is $4\pi r^2 p \,.\, \delta r$. Equating the two expressions obtained for this work, we have

$$4\pi r^2 p \,.\, \delta r = 8\pi\gamma r \,.\, \delta r,$$

or

$$p = \frac{2\gamma}{r}.$$

If the bubble had been a soap bubble this excess pressure would have been $\frac{4\gamma}{r}$, for a soap film has a double surface.

The fact that the pressure inside a soap bubble diminishes as the radius increases is shown by the following experiment. Two brass cups, X and Y, Fig. 5·8 (*b*), about 2 cm. in diameter and 1 cm. long, are connected to stop-cocks A, B, and C as shown. The open ends of X and Y are immersed in a soap solution and soap bubbles differing considerably in diameter blown. B is open and C closed while the larger bubble is being formed, and vice versa. A is then closed and the two bubbles placed in communication with each other by opening the stop-cocks B and C. Air passes from the smaller bubble into the larger one, causing the latter to expand and the former to shrink. This process continues until the radius of curvature of the larger bubble is equal to the radius of curvature of the soap film which finally protrudes below the open end of Y and which is a portion of a spherical surface—see the dotted outlines on the diagram. After a time the thickness of the walls of the large bubble become so thin that it bursts : the film remaining on Y at once becomes flat, and after some time very thin and finally breaks.

Pressure Difference across a Cylindrical Surface.—Let us now assume that Fig. 5·8 (*a*) represents the cross-section of a cylindrical bubble. Since it is difficult to produce such a bubble in a liquid we will assume that it consists of a soap film having *two* surfaces. Consider a length l of this cylinder. When r becomes $r + \delta r$, as before, the increase in area is $2[2\pi(r + \delta r - r)l]$. Let thermal energy be supplied to the film so that its temperature assumes its original value. The increase in the free surface energy

is $4\pi l\gamma \,.\, \delta r$. Now the work done, due to the pressure difference p, is $2\pi r l p \,.\, \delta r$. Equating these two quantities we have $p = \frac{2\gamma}{r}$. When there is only one cylindrical surface the excess pressure is $\frac{\gamma}{r}$.

Angle of Contact.—If a piece of clean glass is inserted into water so that it is in a vertical position, it will be found that the liquid near the glass has been drawn some distance beyond the level of

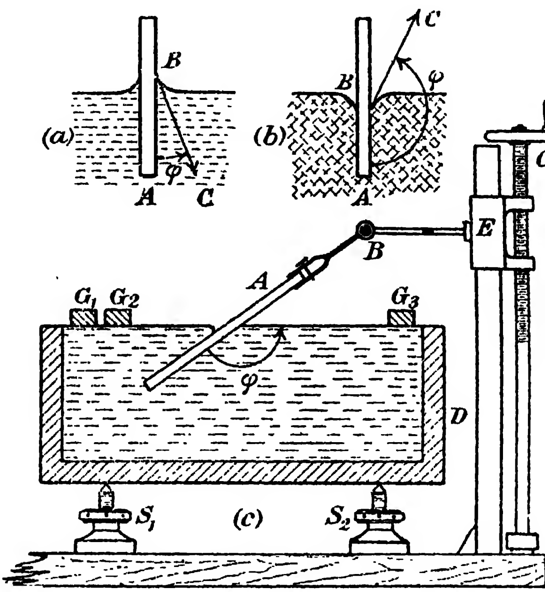


FIG. 5·9.—Angles of Contact and their Measurement.

the rest of the water. The $\widehat{ABC}$, Fig. 5·9 (*a*), i.e. the angle between the solid surface in the water and the tangent to the water surface where it meets the glass, is called the angle of contact for a water-glass interface. For water in contact with glass this angle is very small, whilst for benzol in contact with glass it is zero.

When the above experiment is repeated with mercury the liquid near the glass is depressed below the general level of the mercury surface. The angle of contact is again ABC, Fig. 5·9 (*b*), but it is now quite large (approximately 135°). It should be noted that, although the surface tensions of two liquids may be equal, they may not exhibit the same capillary phenomena, for their angles of contact with a given material may be different.

The effect of the angle of contact on the shape of a small quan-

tity of liquid placed on a flat surface is easily shown as follows :—Water placed on a clean glass surface spreads itself over the glass, but if water is similarly placed on a greased plate it remains as a 'drop.' Traces of dirt or grease alter the angle of contact very considerably ; that is the reason why rain water persists as a drop when it alights on a window-pane, for such a piece of glass is never chemically clean.

To determine the angle of contact between water and glass coated with paraffin wax, N. K. ADAM used an apparatus similar to that shown in Fig. 5·9 (c). A is a section of the plate at right angles to its faces. It is held in a clamp which may be rotated about a horizontal axis through B. The clamp may be moved vertically by means of the screw C, and the carriage E which it operates.

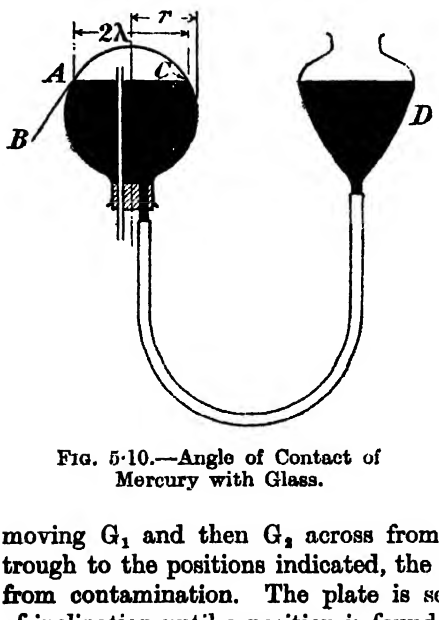


FIG. 5·10.—Angle of Contact of Mercury with Glass.

D is a glass trough, coated inside with paraffin wax so that it may be filled with water above the level of its sides which have been ground flat on the top. This surface is made horizontal with the aid of the screws S_1 and S_2. G_1, G_2 and G_3 are rectangular pieces of glass coated with wax and resting on the sides of the trough, and in contact with the liquid. By moving G_1 and then G_2 across from the right-hand side of the trough to the positions indicated, the surface of the liquid is freed from contamination. The plate is set in turn at various angles of inclination until a position is found for which the water-surface on one side of the plate remains undistorted right up to the line of contact with the solid. If ϕ is the angle between the trace of the plate and the undistorted surface of the water (as measured with the aid of a protractor), then ϕ is the angle of contact required.

In actual practice it is found that ϕ depends on whether the plate A is being pushed into the water or raised. This effect is easily observed by using the rack and pinion to impart the necessary vertical motions to the plate, and the corresponding angles of contact measured in the usual way. If ϕ_1 and ϕ_2 are the 'advancing' and 'receding' angles of contact, it may be shown that $\phi = \frac{1}{2}(\phi_1 + \phi_2)$.

An interesting method for investigating the angle of contact between mercury and glass is as follows :—The level of some mercury in an inverted spherical flask is adjusted by raising or lowering the reservoir D, Fig. 5·10, until the mercury surface in the flask is plane at points where it meets the glass. The angle BAC $= \phi$ is the required angle of contact. If 2λ is the length AC, and r the radius of the flask, $\phi = \sin^{-1}\frac{\lambda}{r}$: it must be remembered that $\frac{\pi}{2} < \phi < \pi$.

Liquid in Contact with a Solid.—We now have to account for the fact that the surface of a liquid near its place of contact with a solid is, in general, curved, even when gravity is the only external force acting throughout the mass of the liquid. Let ABC, Fig. 5·11, be the surface of the liquid. Consider the forces acting on a molecule M in the surface of the liquid and near to the solid D. They are :—

(i) its weight acting vertically downwards;

(ii) the attraction of the solid on M, the direction of which will be along that normal to the surface of the solid which passes through M (since M is very close to the solid);

(iii) the force arising from the attraction of neighbouring liquid molecules. This will be directed towards the interior of the liquid.

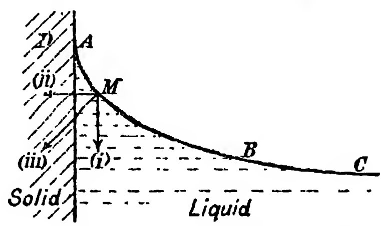


FIG. 5·11.—Liquid in Contact with a Solid.

Now the resultant force exerted on a molecule in an ideal liquid at its free surface must be normal to the surface. Hence the normal to the liquid surface at M will be determined by the resultant of the above three forces. In general, this resultant does not act along (i), i.e. the surface of the liquid at M is not horizontal.

For a molecule near C, a point at a considerable distance from the solid, the only finite forces are (i) and (iii) and these then act vertically downwards, i.e. the surface is flat.

For molecules at B, for example, there is a finite force (ii) but less than the force (ii) on M; in consequence, the surface is more nearly flat.

The Rise of a Liquid in a Capillary Tube.—For the sake of simplicity we shall first assume that the angle of contact is zero. Let AC, Fig. 5·12 (a), be the surface of a liquid in a capillary tube of radius r. We assume that AC is part of a sphere of radius r. The pressure over the curved surface is everywhere atmospheric. At B, a point just below the surface and therefore in the liquid,

the pressure is less than atmospheric by an amount $\frac{2\gamma}{r}$ [cf. p. 116]. At D, a point below B and lying in the same horizontal plane as the surface of the liquid outside the tube, the pressure is atmospheric. Now the difference in pressure between the two points B and D is equal to the pressure exerted by a column of liquid of height DB$=h$ (say). If ρ is the density of the liquid, this difference is $g\rho h$. The pressure at B is therefore less than atmospheric by this amount. But it has already been shown that this difference is $\frac{2\gamma}{r}$. We therefore have

$$\frac{2\gamma}{r} = g\rho h.$$

Now suppose that the angle of contact between the liquid and the material of the tube is ϕ—cf. Fig. 5·12 (b). Let R be the radius of curvature of the liquid surface at its lowest point—if the bore of

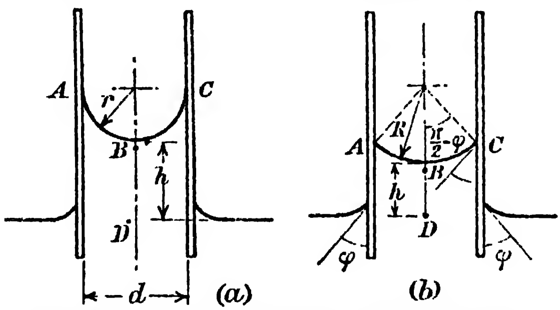


FIG. 5·12.—Rise of Liquids in Capillary Tubes.

the capillary is small, R is constant at all points on the liquid surface. Then, as before, if Π is the atmospheric pressure,

$$\text{Pressure at B} = \Pi - \frac{2\gamma}{\text{R}}.$$

But pressure at D $= \Pi =$ pressure at B $+ g\rho h$.

$$\therefore \frac{2\gamma}{\text{R}} = g\rho h.$$

But $r = \text{R} \cos \phi$; therefore $\frac{2\gamma \cos \phi}{r} = g\rho h$.

It should be mentioned, perhaps, that if ϕ is finite, values of the surface tension of a liquid deduced from measurement of its rise in capillary tubes are unreliable, since the magnitude of ϕ is always uncertain; moreover, ϕ varies considerably with the degree of contamination of the surfaces in contact. The above theory is necessary, however, for academic purposes.

The Rise of a Liquid between Vertical Plates.—(*a*) ***Parallel plates.*** To calculate the amount of this rise we may use Fig. 5·12 (*a*). Let the vertical lines in that diagram now represent sections of the two parallel plates at distance d apart. We assume AC to be a section of a cylindrical surface of diameter d so that the pressure at B is less than atmospheric by an amount $\frac{\gamma}{r}$, or $\frac{2\gamma}{d}$, since $d = 2r$. Proceeding as before we obtain (if the contact angle is zero),

$$\frac{2\gamma}{d} = g\rho h.$$

(*b*) ***Inclined Plates.***—Fig. 5·13 represents two vertical glass plates, AOB and OAD, inclined to one another at a small angle θ. When these are inserted in a liquid the latter rises between the plates. To determine the shape of the curve in which AOC, the vertical plane through OA and bisecting the angle θ, i.e. the plane of co-ordinates, intersects the liquid surface, consider an element PQR of the surface at right angles to the intersection of the liquid surface with the plane AOC. Let (x,y) be the co-ordinates, referred to OC and OA as axes, of Q the middle point of the element PQR. [$P_1Q_1R_1$ is another such element. Notice that the projections p, q, and r of the points P, Q, and R respectively, on the horizontal plane through Ox do not lie in a straight line.] Then if the liquid wets the glass, the surface at PQR is part of that of a cylinder whose diameter is equal to the distance between the plates at Q. This distance is $x\theta$, since θ is small.

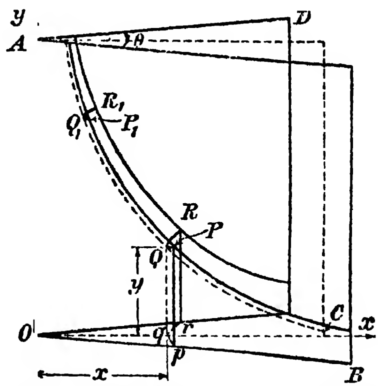


FIG. 5·13.—Rise of a Liquid between inclined Vertical Plates (end effect neglected).

The height y to which the liquid rises is therefore given by

$$g\rho y = \frac{\gamma}{\frac{1}{2}x\theta},$$

i.e. $xy = 2\gamma/g\rho\theta = \text{constant}$. The surface is therefore part of a hyperbola, whose asymptotes are the axes of co-ordinates.

Experimental Determination of Surface Tension.—(*a*) ***Rise in a Capillary Tube Method.*** Select a piece of glass tubing about 0·4 cm. diameter and heat it in a bunsen flame, rotating the tube all the time. When the glass begins to soften, apply a gentle pressure along its length so that the walls of the tube thicken.

Then remove the glass from the flame and *slowly* pull the ends apart. The capillary tube thus constructed is clean, a condition which is absolutely essential if a reliable value for γ is to be obtained. When the tube is cold select a length from the centre of the drawn-out portion and attach to it a very thin glass rod, R, drawn out to a point and bent twice at right angles as in Fig. 5·14. Bands B_1 and B_2 cut from a length of rubber tubing enable this rod to be attached to the tube easily.

Now clamp the capillary A in a vertical position and place the liquid whose surface tension is to be measured below the tube so that the latter is immersed to a greater depth than that at which it is to be used and then raise it slightly. If the liquid falls back readily as the tube is raised we may assume that the tube and liquid are not contaminated. Continue to raise the tube until the end of the rod is just about to break through the liquid surface. To measure the height of the liquid in the capillary a vernier microscope, M, should be used. The microscope is focussed on the lowest point of the liquid surface in the capillary and the reading on its scale observed. The vessel containing the liquid is then removed, care being taken to see that the rod is not disturbed. The microscope is then focussed on the end of the rod and the reading noted. The difference between these readings gives the height of the liquid in the capillary. These observations should be repeated. The tube is then broken at the point corresponding to the top of the meniscus and the radius found with the aid of a vernier microscope. To do this several readings of two diameters mutually at right angles are made. If the mean values of each set are equal to within about 5 per cent. the mean value can be taken as a measure of r. If the discrepancy is greater than this the tube should be rejected and another one constructed. It often saves much time if the mean diameters of the two ends of the tube are measured before commencing the experiment. If these are circular the chances will be that the rest of the tube will have a circular section. But these values must not be used in calculating γ since it is the radius at the point B, Fig. 5·12, which

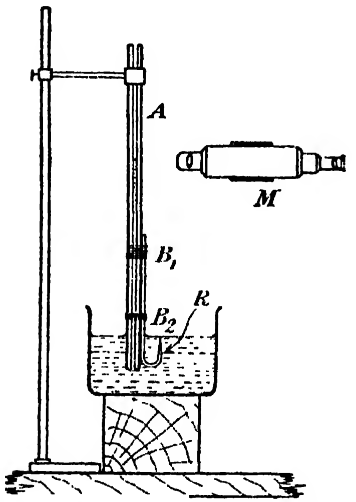


FIG. 5·14.—Measurement of Surface Tension by Rise of Liquid in a Capillary Tube.

determines the pressure change in crossing the surface of the liquid. The value of the surface tension may then be calculated from the formula already proved.

[At this point it is convenient to ask ourselves what would happen if a tube of radius r and length less than h, where h is given by $2\gamma = g\rho hr$, were dipped in a liquid of surface tension γ and density ρ. Usually, i.e. when the length of the tube is greater than h, it is the height of the liquid in the tube which adjusts itself until the equation is satisfied. When this is no longer possible, as in the problem now contemplated, the only quantity in the above equation which is a variable is r. The liquid therefore rises to the top of the tube and there forms a surface which is concave upwards and whose radius is greater than r. Its value r_1 is given by $h_1r_1 = hr$, where h_1 is the height of the liquid in the capillary.]

Note on Comparing Experimentally the Surface Tensions of Two Liquids.—If the 'rise in a capillary tube' method is adopted it is not necessary to determine the radius of the tube if the tube is arranged so that the liquid meniscus stands in turn at the same position in the tube when the heights to which the liquids rise are determined. Then

$$\gamma_1 = \tfrac{1}{2}g\rho_1h_1r, \text{ and } \gamma_2 = \tfrac{1}{2}g\rho_2h_2r.$$

$$\therefore \frac{\gamma_1}{\gamma_2} = \frac{\rho_1h_1}{\rho_2h_2},$$

and $\frac{\rho_1}{\rho_2}$ may be determined directly by means of Hare's apparatus.

(*b*) ***Jaeger's Method or the Method of Maximum Bubble Pressure.***—This is based on the fact that the excess pressure inside a spherical bubble of air inside a liquid is $\frac{2\gamma}{r}$ where r is the radius of the bubble.

The experiment consists essentially in determining the maximum pressure required to produce an air bubble at the end of a vertical capillary tube immersed in the liquid whose surface tension is being determined. A capillary tube about 0·05 cm. in diameter is constructed as in (*a*). This is placed vertically downwards in a vessel, A, Fig. 5.15 (*a*), containing the liquid whose surface tension is required. It is connected to a manometer, C, containing xylol, and also to a Woulf's bottle, D, fitted with a dropping funnel, B. Mercury (or water) is placed in B and permitted to run slowly into D. A difference of pressure between the inside and the outside of the apparatus is at once shown if the apparatus is air-tight. When the pressure in D reaches a certain value bubbles appear in A. These should be formed singly and at the rate of about one in ten

seconds. The first condition is obtained by reducing the volume of air in the apparatus so that when one bubble breaks away from the end of the capillary tube, the pressure inside the apparatus is reduced to such a value that it is less than the maximum pressure required to blow the bubble; the second condition is obtained by adjusting the rate at which liquid flows into D. The maximum height h of the manometer is recorded. If ρ is the density of the liquid in the gauge, the pressure recorded by it is $g\rho h$, where g is the acceleration due to gravity. But this pressure difference is not entirely due to the effects of surface tension, for part is attributable to the pressure due to the fact that the orifice of the capillary is at a depth d below the surface of the liquid. If σ is the density of this liquid,

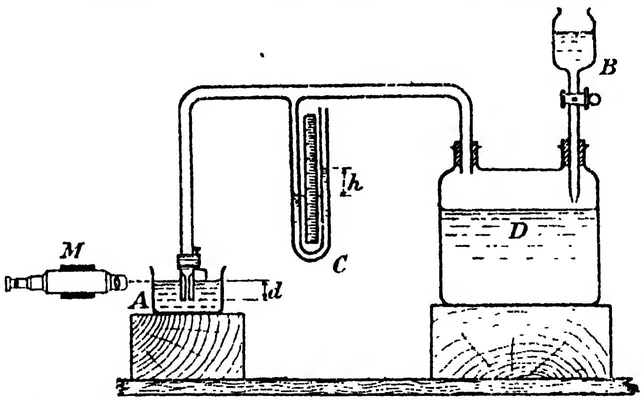


FIG. 5·15 (a).—Apparatus for Determining Surface Tension of a Liquid. [The microscope M is only used when the vessel A has been removed, for otherwise refraction at the curved surface of this vessel would vitiate the observations.]

this pressure amounts to $g\sigma d$, so that the pressure difference directly attributable to surface tension is $g[\rho h - \sigma d]$. We therefore have

$$\frac{2\gamma}{r} = g(\rho h - \sigma d).$$

Hence γ may be calculated when the other variables in this equation are known.

To discover the reason why the value of r used in the above equation is equal to the radius of the capillary tube at its lower end, let us suppose that the tube is uniform in diameter and that the pressure inside the apparatus is such that the centre of the hemispherical liquid surface is at C_1—cf. Fig. 5·15 (b). We are justified in assuming that this surface is part of a sphere if the radius of the capillary is not large, and the angle of contact between the liquid and the tube is zero. Suppose that the pressure inside the apparatus is increased so that the centre of

the surface is at C_2, the radius still being r, but that if the surface is forced down beyond this position its radius increases. When C_3 is the centre, let the radius be $(r + \delta r)$. The pressure difference across the surface is then less and the bubble grows since the pressure inside the apparatus is too great for the surface to be in equilibrium. Thus a bubble of air escapes, and the liquid surface will lie entirely above C_2, if the removal of one bubble is sufficient to reduce the pressure inside the apparatus below the maximum pressure necessary to cause a bubble to escape from the tube. If not, several bubbles will escape.

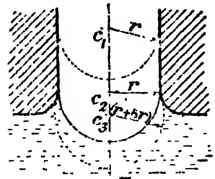


FIG. 5·15 (*b*).—Formation of a Bubble at the end of a Capillary Tube (greatly enlarged).

The great advantages of this method are that it may be applied to determine the surface tension of a molten metal, or to investigate how the surface tension of a liquid varies with temperature, or how that of a solution varies with the concentration of the dissolved substance. The method is particularly suited for such determinations as the two last, since it is not necessary to know the radius of the capillary tube. Also, since a new surface is continually being formed in the liquid the effects of contamination are reduced to a minimum, and finally the radius r can be determined before observations are made [cf. method (*a*)].

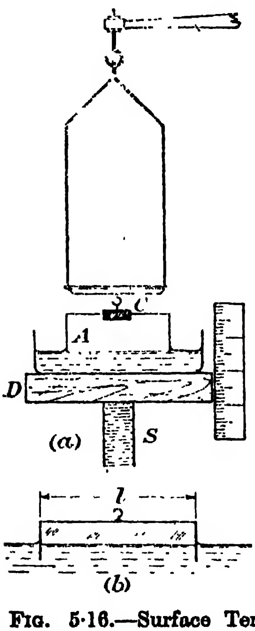


FIG. 5·16.—Surface Tension by ordinary Balance Methods.

Unfortunately certain difficulties arise when an absolute determination of the surface tension of a liquid is being made by this method. One is seldom quite sure whether or not the size of the bubbles, when the excess pressure inside the bubble is a maximum, is controlled by the internal or the external radius of the tube. If these radii differ considerably and the surface tension is known at least approximately, simple substitution of these values in the appropriate equation reveals the correct one. In addition, although for many years it has been maintained that the method gives results which are independent of the angle of contact of the liquid with the material of the tube, PORTER has recently shown, at least for angles of contact greater than

$\pi/2$, that when the external radius is the determining one, the calculation does not involve the angle of contact, but that it is quite otherwise when the excess pressure is determined by the internal radius. Porter also remarks that it is a matter of some surprise that the belief that the results were always independent of the angle of contact should ever have gained credence, although that belief is generally held.

(*c*) ***Ordinary Balance Method.***—The surface tension of a liquid which wets glass may be determined as follows. A glass plate, A, Fig. 5·16 (*a*), (a microscope slide) is supported by means of a metal clip, C, from below the pan of a balance—the lower edge of the slide is made horizontal. The vessel, D, containing the liquid is placed on a small table below the slide. The table may be raised by means of a screw, S. The balance is equilibrated and left free to swing. The adjustable table is then screwed up till the liquid *just* touches the lower edge of the plate. This is shown by a sharp jerk of the pointer as the microscope slide is pulled down by surface tension. Masses are then added to the other pan of the balance until the slide is withdrawn from the liquid. Since the lower edge of the slide had been in the general level of the liquid surface there is no correction for buoyancy. If l is the length and t the thickness of the slide at its lower edge, the force due to surface tension acting on it is $2(l + t)\gamma$. This is equal to mg, where m is the mass added to the pan to restore equilibrium. Hence γ may be determined.

An alternative method is as follows. Having screwed up the adjustable table till the pointer jerks, observe the position of the table (suitable scales may be arranged as on a spherometer). Instead of restoring equilibrium as above, the table is screwed up through a distance h until the pointer is back at zero. Then the buoyancy force just balances the force due to surface tension, and if the vessel containing the liquid has a large surface area, so that h will be also the depth of immersion of the slide, then

$$2(l+t)\gamma = lth\rho g,$$

where ρ is the density of the liquid.

It must be noted that this method only yields accurate results if the liquid completely fills the containing vessel so that the surface of the liquid may be cleaned with the aid of waxed pieces of glass, as described on p. 118.

Soap Solutions.[1]—The plate method described above may

[1] Prof. Boys recommends the following soap solution. To a litre of distilled water contained in a well-stoppered bottle add 25 gm. of sodium oleate, and let it stand for 24 hours. Then add about 300 cm.^{3} of glycerol, shake well, and allow to stand for a week. By means of a siphon remove the clear

easily be adapted to determine the surface tension of a soap solution. A glass or wire frame, as shown in Fig. 5·16 (*b*), is made and is supported from below one pan of a balance, and arranged that when the balance is equilibrated, the horizontal portion of the frame is about 0·5 cm. above the general surface of the liquid. The frame is then immersed completely and extra masses, m, added to the right-hand balance pan until the frame is in the same relative position as before. If l is the length of the horizontal portion, the weight of the film being negligible,

$$2\gamma l = mg.$$

[This method may be used for liquids such as water, the horizontal portion of the frame then being nearer to the general surface of the liquid.]

Drops and their Formation.—Suppose that a glass tube about 2 mm. in diameter has been connected to a wide tube by means of rubber tubing and a narrow capillary glass tube, and the whole filled with a liquid—say, water. The capillary tube is merely to control almost entirely the rate at which the liquid escapes when the apparatus is held in a vertical position with the narrow tube pointing downwards. If the water leaving the tube is carefully watched it will be seen to assume, in turn, shapes whose outlines are shown in Fig. 5·17 (*a*) and (*b*). As the drop continues to grow a waist is formed—the drop is then about to break away —cf. Fig. 5·17 (*c*). When this occurs the water comprising the neck will form a small sphere following the larger drop. It is known as ***Plateau's spherule*** —cf. Fig. 5·17 (*c*).

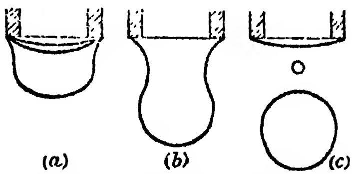


FIG. 5·17.—Drops and their Formation. (*a*) Early Stages in the formation of a Drop. (*b*) Drop showing the formation of a Waist. (*c*) A large Drop and Plateau's Spherule.

To observe more easily the formation of a drop of liquid it is necessary to diminish the effective pull of gravity on the drop. This was done in a very striking way by DARLING. At temperatures above 80° C. the density of aniline is less than that of water at the same temperature, whereas the reverse is true at lower temperatures. Moreover, aniline and water are immiscible. Suppose, therefore, that a large tall beaker is nearly filled with water and a quantity of aniline (about 100 cm.^{3}) added. This collects at the

liquid, leaving the scum behind. Add two or three drops of liquid ammonia to the solution and store in a dark cupboard. The solution must not be warmed or filtered.

bottom of the beaker. A bunsen burner is then placed below the beaker: when the aniline assumes a temperature of about 80° C. it ascends to the top of the water and collects there in the form of a pendant drop. The rate of supply of heat is diminished and the aniline cools: a large drop about 3 cm. in diameter begins to form. The drop then has a distinct neck which gradually becomes more thin. Finally, two constrictions are formed, and a large drop of aniline, followed by Plateau's spherule, falls to the bottom of the beaker. The process is then repeated.

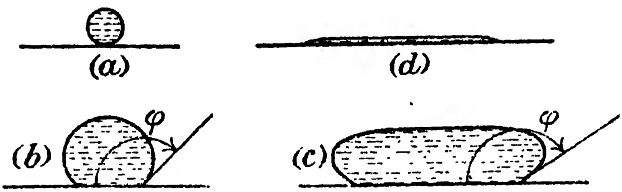


FIG. 5·18.—Wetting of a Surface by a Liquid.

Drops of liquid not wetting the surface with which they are in contact—say, mercury—are, when small, truly spherical—cf. Fig. 5·18 (*a*). As the drop grows, or if several small ones coalesce, it loses its sphericity—cf. Fig. 5·18 (*b*)—while a very large drop is perfectly flat except near its edges, i.e. it assumes the shape shown in Fig. 5·18 (*c*).

If the liquid wets the surface, drops do not form, and the liquid spreads itself over the surface—Fig. 5·18 (*d*).

A Liquid Drop between Plates.—When two clean pieces of plate glass are placed face to face no difficulty is experienced in separating them, but a considerable force is necessary to pull them away when a small drop of water, for example, is placed between them and the plates are very close together. If d is the distance apart of the plates and the area wetted is large, then we may regard any small element of the liquid surface as part of a cylindrical one with a radius of curvature $0{\cdot}5d$. Consequently the pressure in the water is less than that outside by an amount $\frac{2\gamma}{d}$. If A is the area of each plate which is wetted, the total force pulling the plates together is $\frac{2\gamma}{d}.A$.

On the other hand, when a small drop of mercury is placed between two plane surfaces, considerable force must be applied to the plates in order to flatten the drop to any extent. If ϕ is the angle of contact between mercury and glass, the radius of curva-

ture, R, of the mercury surface normal to the plane of the plates is given by

$$\frac{\frac{d}{2}}{R} = -\cos\phi.$$

$\therefore$ Pressure in the mercury is greater than atmospheric by

$$\frac{\gamma}{R} = -\frac{2\gamma\cos\phi}{d}.$$

Since $\frac{\pi}{2} < \phi < \pi$, the above expression is positive. The force to be applied is therefore $\frac{2\gamma A\cos(\pi - \phi)}{d}$.

Surface Films on Water.—Many pure substances of a fatty nature, when placed on a clean surface of water, spread themselves out to form an exceedingly thin surface layer. It can be shown that the thinnest film which can be formed on water is one molecule in thickness, each molecule of the oil being in direct contact with the surface of the water.

Some extremely interesting conclusions have been drawn from the study of these layers, for their simple structure makes them peculiarly suitable for investigating the properties of the molecules themselves. It is found that these films can exist in three forms, corresponding to the solid, liquid and gaseous states of matter. In the 'gaseous' state of the films, the molecules move about individually and separately in the surface, exerting an outward spreading force on the boundary of the surface, in much the same manner as a gas exerts a pressure on the walls of the vessel containing it, or a dissolved substance exerts an osmotic pressure on a semi-permeable membrane. In the 'solid' and 'liquid' states of the film the molecules adhere into compact, coherent masses, in which they are often just as closely packed as in solids or liquids in bulk.

In these coherent films the cross-sectional area of the individual molecules has been measured by measuring the area of the film composed of a known number of molecules, as calculated from the mass of the film, i.e. the mass of the drop of substance placed on the water surface, and the mass of one of its molecules. The results of such measurements show that the molecules actually have the shapes which have been indicated for about three-quarters of a century by the structural formulæ of organic chemistry. It is found, for instance, that the molecule of stearic acid, $C_{17}H_{35}COOH$, is just about five times as long as it is thick; that the end group (COOH), under certain circumstances, is slightly thicker than the rest of the molecule; and that usually the molecules pack into a coherent layer,

standing nearly vertical with the COOH groups directed towards the water. Many other coherent films, though not all, have the same vertical disposition of the molecules. In the 'gaseous' films, when the molecules do not cohere, they lie flat upon the surface of the water.

Viscosity.—Whenever relative motion exists between the different layers into which we may imagine a liquid is divided, forces are called into play tending to retard the more rapidly moving layers and to accelerate those which are moving more slowly. Similar forces, although much smaller, arise when a gas moves in the same way. To obtain a more definite idea of these forces let us consider Fig. 5·19 (*a*). In this xOy represents the boundary between a fluid and a solid over which the former is flowing. At this boundary it will be assumed that the fluid is at rest, and that all the molecules in a plane parallel to xOy have a resultant motion

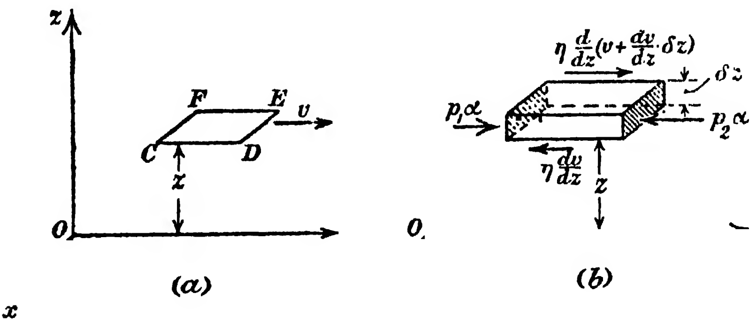


FIG. 5·19.—Coefficient of Viscosity.

(mass-velocity of the fluid) which is parallel to the above reference plane and which increases with the distance of the layer from that plane. Let CDEF be an area of magnitude S at distance z from the reference plane. Then the molecules immediately above this plane tend to accelerate the molecules in it, while the molecules in the layer immediately below tend to retard them. In this way each stratum of fluid will exert on the one next to it a tangential traction, opposing the relative motion between the two layers. If F is the magnitude of this tangential force, the force acting on unit area of CE is $\frac{F}{S}$: this is the tangential stress due to viscosity in the fluid. We assume that the magnitude of this stress is directly proportional to the difference in velocity between the layers immediately above and below the plane considered, divided by their distance apart. This latter quantity is termed the ***velocity-***

gradient in the fluid. It is denoted by $\frac{dv}{dz}$, where v is the mass-velocity of the fluid at a height z above the reference plane. We may therefore write

$$\frac{F}{S} = \eta\frac{dv}{dz},$$

where η is a constant called the ***coefficient of viscosity*** of the fluid. It depends upon the nature of the liquid and its temperature. [Notice the similarity between this definition and those of diffusion and of thermal conductivity.]

The value of η in C.G.S. units is expressed in dynes per square centimetre per unit velocity gradient, i.e. gm. cm.$^{-1}$ sec.$^{-1}$ This unit is often called the '***poise***' in honour of POISEUILLE.

The above equation cannot be verified directly, but calculations based on it are in strict accord with experiment so that we do not hesitate to accept the above equation as a complete statement of the laws of viscosity.

To determine the relation between the viscous forces in a fluid and the pressure differences in it, consider the volume of fluid lying between planes at heights z and $z + \delta z$ above xOy—cf. Fig. 5·20 (b). Let the area of the faces parallel to xOy be unity, and let α be the cross-sectional area of the element in a direction normal to Oy.

Then the forces due to viscosity acting on the lower and upper faces are

$$\eta\frac{dv}{dz} \text{ and } \eta\frac{d}{dz}\left(v + \frac{dv}{dz}\cdot\delta z\right),$$

their lines of action being parallel to yO and to Oy respectively. Let p_1 and p_2 be the pressures at points on the two ends of the prism, $p_1 > p_2$; the forces are $p_1\alpha$ and $p_2\alpha$ as indicated. Since the fluid is moving without acceleration, the total force on the element considered must be zero. Hence

$$\eta\frac{d}{dz}\left(v + \frac{dv}{dz}\cdot\delta z\right) - \eta\frac{dv}{dz} + p_1\alpha - p_2\alpha = 0.$$

$$\therefore\ \eta\frac{d^2v}{dz^2}\cdot\delta z = (-p_1 + p_2)\alpha.$$

Experimental Determination of Viscosity.—*Method i:* To determine the viscosity of water the apparatus shown in Fig. 5·20 may be used. It consists of a tall metal cylinder furnished with an overflow pipe DC. A capillary tube of known length, l, and radius, r, is placed in a horizontal position and connected to the cylinder. Water enters along the inlet tube as shown, any excess being carried away along DC. Attached to the exit end of the capillary is a glass tube bent in the manner indicated. The pressure difference between the ends of the capillary is proportional to the

vertical distance between the levels A and B. This may be determined with the aid of the scale in mm. and a U-tube filled with water and placed as shown so that the levels at A and C are the same. If the water is allowed to flow along the tube, as each drop breaks away from E the water level at B changes—an effect due to the changes in pressure at E as the drops alter in shape. This disturbing factor may be avoided if a small clean glass rod is placed in contact with the liquid. The liquid then leaves the tube in a trickle and the level at B is constant.

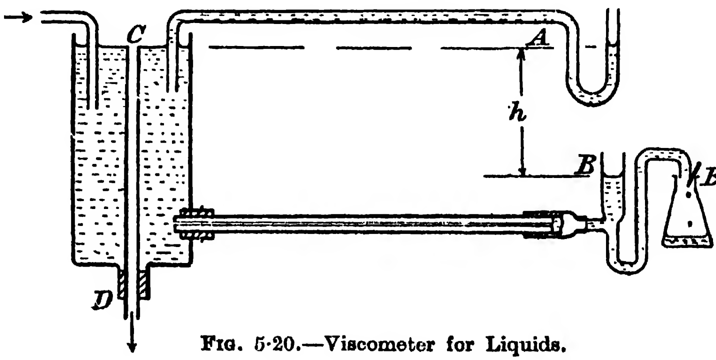


FIG. 5·20.—Viscometer for Liquids.

It may be shown, by reasoning beyond the scope of this book, that v, the volume of liquid emerging in t seconds from the tube, is given by

$$v = \frac{\pi r^4 pt}{8l\eta},$$

where η is the coefficient of viscosity, p the pressure difference between the ends of the capillary—it is $g\rho h$, where $h = AB$, ρ is the density of the water, and g is the acceleration due to gravity. It must be pointed out that the formula is true only for narrow tubes in which the velocity of the liquid is not so great that the flow becomes turbulent.

Stokes' Law.—When a sphere falls vertically downwards through a viscous medium, the layers of liquid adjacent to the sphere tend to move with a velocity equal to that of the sphere. At a great distance from the sphere the liquid is at rest. Consequently there must be relative motion between the different layers of the liquid and the motion of the sphere will depend on the viscosity of the medium. If the sphere is small it is found that it soon acquires a constant velocity, i.e. the pull due to gravity on the sphere is balanced by the upthrust of the liquid on it and the force arising from its motion through the viscous medium.

This vertical force, F, will depend on η, the viscosity of the medium,

a, the radius of the sphere, and v the constant or terminal velocity acquired by the sphere. Thus

$$F = \kappa a^{\alpha}\eta^{\beta}v^{\gamma}$$

where κ is a constant, and α, β, and γ are the appropriate dimensional coefficients. In addition to the dimensions of F, a, g, and v, which are already known, we require those of η. Now

$$\frac{[\text{force}]}{[\text{area}]} = [\eta]\left[\frac{dv}{dx}\right].$$

Hence $$\frac{[MLT^{-2}]}{[L^2]} = [\eta]\frac{[LT^{-1}]}{[L]},$$

so that $[\eta] = [M][L]^{-1}[T]^{-1}$.
We therefore have

$$[MLT^{-2}] = [L]^{\alpha}\,[ML^{-1}T^{-1}]^{\beta}\,[LT^{-1}]^{\gamma}.$$

Equating like exponents, we have

$$\beta = 1,\ \alpha - \beta + \gamma = 1,\ \beta + \gamma = 2.$$
$$\therefore\ \gamma = 1,\ \alpha = 1.$$
$$\therefore\ F = \kappa a\eta v,$$

and it can be shown that $\kappa = 6\pi$, i.e. $F = 6\pi a\eta v$. This expression was first obtained by STOKES, and is known as Stokes' law for the force acting on a sphere falling under gravity through a viscous medium.

The Viscosity of Oils.—Suppose that ρ is the density of the material of the sphere, σ that of the liquid. Since

Weight of sphere = upthrust due to liquid displaced + force due to the motion of the sphere, we have,

$$\tfrac{4}{3}\pi a^3\rho g = \tfrac{4}{3}\pi a^3\sigma g + 6\pi a\eta v.$$

$$\therefore\ \eta = \tfrac{2}{9}a^2g\frac{(\rho - \sigma)}{v}.$$

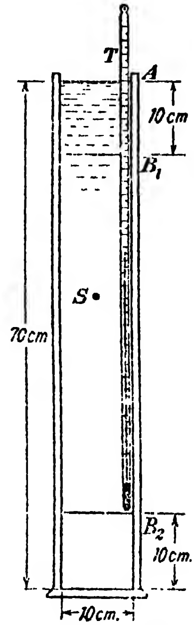


FIG. 5·21.—Viscosity of Oils—Stokes' Method.

The above expression shows that if the velocity of fall of a sphere through a viscous medium can be measured, we have a means of determining the coefficient of viscosity of the medium.

Let us suppose that glycerol is the liquid whose viscosity is to be determined. This is placed in a glass cylinder, A, Fig. 5·21, about 70 cm. long and 10 cm. wide. Spheres of known diameter are dropped into the liquid and the terminal velocity for each sphere is deduced from observations on the time required for the sphere to travel between two fiducial marks. Now the liquid is limited by the walls of the vessel and has a finite depth. The conditions stipulated by the above theory are therefore not fulfilled. It may be shown, however, that if the sphere falls between two fiducial marks B_1 and B_2 (10 cm. from the top and bottom of the liquid respectively), then the motion is uniform. Further, if the diameter of the sphere does not exceed 0·2 cm. and a vessel 10 cm. wide is used, no correction is necessary for the effect of the walls of the vessel. If λ is the distance between the fiducial marks, and t the time of transit,

$$\eta = \tfrac{2}{9}a^2g\left(\frac{\rho - \sigma}{\lambda}\right)t,$$

so that a^2t is constant for a given liquid at a constant temperature. If therefore a^2 is plotted against $1/t$ a straight line should be obtained if the conditions of the theory have been satisfactorily fulfilled. From the slope of the line η may be deduced.

Since the viscosity of an oil changes very rapidly with temperature it is advisable to measure and record the latter to within 0·1° C.—T is a thermometer—and, having measured the diameters of all the spheres to be used in some definite experiment, to carry out the fall experiments one after the other as quickly as possible. The spheres must fall centrally down the tube A, so that their fall shall not be affected by the walls of the tube.

EXAMPLES V

1.—How would you proceed to determine the osmotic pressure of a solution? Give some account of plasmolysis and isotonic solutions.

2.—Describe the apparatus used for preparing colloids by means of hot dialysis.

3.—A liquid whose density is 0·83 gm. cm.$^{-3}$ rises to a height of 8·92 cm. in a tube whose diameter is 0·0168 cm. What is the surface tension of the liquid?

4.—Describe a method of determining the viscosity of an oil. In an experiment with the cup and ball viscometer the time to break away for an oil of known viscosity 6·3 gm. cm.$^{-1}$ sec.$^{-1}$ was 60·7 sec. What is the viscosity of an oil when the time is 26·2 sec.?

5.—Define the terms *surface tension* and *surface energy*. Give the theory of one method of determining the surface tension of a liquid whose angle of contact with glass is zero. How would you demonstrate the existence of surface energy in a liquid film?

6.—What determines whether a liquid will rise or fall in a capillary tube placed with one end below the surface of a liquid? How may the surface tension of a molten metal be determined?

7.—Explain the terms *osmosis* and *osmotic pressure*. Upon what factors does the osmotic pressure of a solution depend?

8.—A glass microscope slide, 10 cm. long and 1 mm. thick, is suspended from one arm of a balance so that its lower edge is horizontal and its plane vertical. The balance is left free and equilibrated. A vessel containing alcohol is placed below the slide and then raised until the alcohol just touches the lower edge of the slide. If a mass of 0·63 grams must be placed in the opposite pan of the balance to restore equilibrium, calculate the surface tension of alcohol.

9.—Define the coefficient of viscosity and describe how you would proceed to compare the viscosities of two liquids—say alcohol and water—at room temperature.

10.—Describe and explain how the surface tension of a liquid may be measured by forcing bubbles of air through it. Discuss whether the result obtained in this way should be the same as that given by the capillary tube method.

11.—Two vertical plates, distance d apart, are immersed in a liquid whose angle of contact with the plates is zero, and whose surface tension is γ. Calculate the height to which the liquid will rise at a point some distance from the edges of the plates.

12.—Discuss the shape of a liquid surface in the space between two vertical plates inclined at a small angle to one another.

13.—Describe and explain what happens when minute camphor particles are scattered on a clean water surface. Why does immersing one's finger in the water modify the effect? A spherical soap bubble of radius 2 cm. is blown in an atmosphere whose pressure is 10^6 dyne. cm.$^{-2}$. If the surface tension of the liquid composing the film is 60 dyne. cm.$^{-1}$, to what pressure must the surrounding atmosphere be brought in order exactly to double the radius of the bubble? Assume no temperature change and no diffusion through the bubble. (N.H.S.C. '29.)

14.—State and give the theory of a method of determining the surface tension of mercury, in which measurement of the angle of contact between mercury and glass can be avoided. (L. '23.)

15.—Define *surface tension* and *angle of contact*. If the surface tension of a liquid having a density of 0·82 gm. cm.$^{-3}$, is 28·3 dyne. cm.$^{-1}$, calculate the height to which the liquid will rise in a glass capillary tube of 0·5 mm. diameter dipped into it, the angle of contact between the liquid and glass being 30°.

16.—A U-tube with vertical limbs is half-filled with liquid. If the diameters of the two limbs are 1 cm. and 0·1 cm. respectively, calculate the difference in height of the liquid in the two limbs if the density of the liquid is 1·27 gm. cm.$^{-3}$ and its surface tension is 45 dyne. cm.$^{-1}$ Assume the angle of contact to be zero.

17.—A capillary tube 0·15 mm. in diameter has its lower end immersed in a liquid whose surface tension is 54 dyne. cm.$^{-1}$ and whose density is 0·86 gm. cm.$^{-3}$ Calculate the height to which the liquid rises, the angle of contact being 28°. Establish the formula used.

CHAPTER VI

ELASTICITY

Strain and Stress.—A system of forces acting on a body may sometimes be such that although there is no motion of the body as a whole yet there may be a relative displacement of its constituent particles causing a change of form or a change in the dimensions of the body. Such a body is said to be ***strained.*** When a body is strained forces are called into play tending to resist the relative displacement of the component particles: the body is then said to be in a ***state of stress.*** There are three types of simple strain and simple stress: (*a*) tensile strain and tensile stress, (*b*) compressive strain and compressive stress, and (*c*) strain and stress caused by shear.

Tensile Strain and Stress.—In Fig. 6·1 (*a*), AB represents a uniform bar of initial length L. When stretching forces FF act upon AB its length increases by an amount l when equilibrium is attained, i.e. the internal forces in the body have reached such a magnitude that a further displacement of the component particles of the body is prevented.

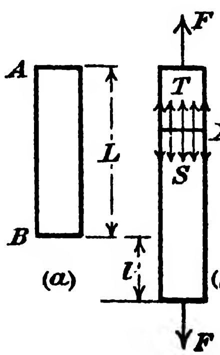


FIG. 6·1.—Tensile Strain and Stress.

The ratio $\frac{l}{L}$ is called the ***tensile strain*** of the body and since both l and L are lengths, this strain, like every other strain, is measured by a mere number.

Taking any arbitrary and imaginary section in the bar normal to its length as at X, Fig. 6·1 (*b*), the internal forces across this section are such that the forces S just balance the force F at A, while the forces T just balance F at B. These internal forces resist the efforts of the forces FF to break the bar: they constitute a ***tensile stress.*** Since these internal forces are distributed over an area the ***stress*** is measured by the force per unit area, so that stresses are expressed in the absolute systems of units either as dyne. cm.$^{-2}$, or as poundal.ft.$^{-2}$ Since the resultant of the

internal forces S or T at X is F, if P is the stress and s the area of cross-section at X, we have $P = \frac{F}{s}$.

Compressive Strain and Stress.—If the forces FF acting on the above body were reversed the length would decrease by an amount l, the body would be subject to a ***compressive strain*** of amount $\frac{l}{L}$, and the stress due to compression would be $\frac{F}{s}$.

Shear Stress and Strain.—A shear stress exists between two parts of a body in contact when each part exerts an equal and opposite force laterally on the other part and in a direction tangential to the surface of contact separating the two parts. Thus, suppose a rivet holds two plates together which sustain a pull F,F, across the section AB, Fig. 6·2. Under these conditions the lower portion of the rivet exerts a force parallel to AB on the upper portion, preventing it from moving to the left: similarly, the upper part exerts a force on the lower. The rivet is said to be

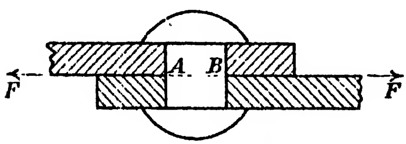


Fig. 6·2.—Shear Stress and Strain.

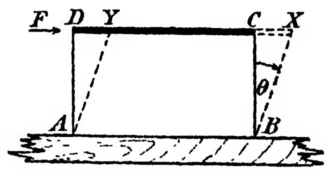


Fig. 6·3.—Shearing Strain.

in a ***state of shear*** across the plane AB, and if s is the area of the section AB, the ***shear stress*** is defined as F/s.

To discover how the strain is measured when a state of shear exists, let us consider ABCD, Fig. 6·3, the cross-section of a block of india-rubber glued to a table along that face of which AB is the trace. Imagine that a piece of sheet brass glued to the upper surface is urged forward by a force F parallel to AB. When equilibrium is reached let the plate be in position XY, i.e. the plate will have suffered a displacement CX with respect to the lower face. The block is now said to be sheared, the amount of the shearing strain being specified by the ratio $\frac{CX}{BC}$, i.e. $\tan\theta$, where θ is the $\widehat{CBX}$. It will be seen that the shearing strain is the ratio of the relative lateral displacement CX of two horizontal layers at distance BC apart to that distance, i.e. it is equal to the numerical value of the relative lateral displacement of two horizontal layers at unit distance apart.

If s is the area of the upper face the shearing stress is $\frac{F}{s}$.

It is important to note the following distinction between strain due to stretching [or compressing] forces and that due to shearing forces, for in the first instance both the volume and shape of the body may alter, whereas in the second it is the shape alone which changes, the volume remaining constant. A particular instance in which a change in volume but no change in shape occurs is when a cube of material which is isotropic, i.e. has properties the same in all directions, is subjected to a uniform pressure.

Complimentary Stresses due to Shear.—Theorem: ***A shear stress in a given direction cannot exist without an equal shear stress existing at right angles to it.*** To prove this, let us consider the rectangular body of sides, a, b, and c, shown in Fig. 6·4. Let F_1, F_1, be the forces tending to displace the upper face with respect to the lower. The area of each of these faces is ab, so that the shear stress is F/ab. Let F_2,F_2 be shearing forces at right-angles to the above. Then the corresponding stress is F_2/ac. For equilibrium, the moment of all the forces about any point in their plane—say A—must be zero, i.e. $F_1 . c = F_2 . b$. Dividing throughout by abc, we have

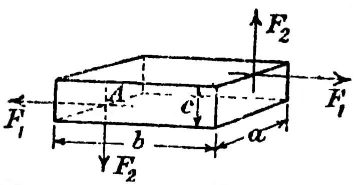


FIG. 6·4.—Shear Stress and Strain

$$\frac{F_1}{ab} = \frac{F_2}{ac},$$

i.e. the stresses are equal.

Elasticity.—When the forces acting on a strained material are removed the body may assume its original form and dimensions. Such a body is said to possess ***elasticity***. Thus a piece of india-rubber is very ***elastic***, while lead and putty are almost ***non-elastic***.

Hooke's law and the Limit of Perfect Elasticity.—We have just defined the terms elastic and non-elastic as if they applied to two essentially different classes of substance. Actually, all bodies are elastic to a certain degree, depending on the magnitude of the applied load. Thus, if lead is subjected to small stretching forces it recovers its original form and size when the forces are removed —i.e. the lead then behaves as an elastic body. On the other hand, lead is non-elastic when the forces are not small and it is said to acquire a ***permanent set***. The limit of stress within which the strain in a given material completely disappears when the stress is removed is called the ***elastic limit***, or ***limit of perfect elasticity***. At stresses below the elastic limit there is a linear relationship between a stress and the corresponding strain: this

fact was discovered in 1679 by HOOKE, a contemporary of BOYLE, and is known as *Hooke's law.*

The existence of the elastic limit is very strikingly shown by the following experiment:—Two long pieces of copper wire of the same diameter are suspended from the ceiling and an electric current passed through one of them so that it just glows in a darkened room. When the wire is cool, pans are attached to each wire and loaded with equal masses which are increased by 250 or 500 gm. at a time. At first the elongation of each wire is of the same order of magnitude and if the loads are removed the wires will resume their original lengths. On increasing the loads further, a stage is soon reached when the wire which has been heated extends very rapidly and when the load is removed it is found to have acquired a permanent set. From this experiment it is clear that the elastic properties of a given material depend on its previous history. The heated wire is in an annealed condition, whereas the other wire which has been manufactured by drawing it through a die [a small hole in a steel plate—called a *Wurtel* plate] is said to have been ***cold worked.*** The effect of cold-working a metal by drawing it through a die, by rolling it in a mill, or by hammering it, is to increase its hardness, to lower its ductility, and to diminish its capacity for resisting mechanical shocks.

Young's Modulus and the Modulus of Rigidity.—The fundamental law relating to elasticity, discovered by HOOKE, is that the strain is proportional to the stress by which it is produced, provided that the elastic limit has not been exceeded. This relationship may be written

$$\text{stress} = k \times \text{strain}$$

where k is a constant in any given instance. This constant is called the ***modulus of elasticity*** and depends upon the nature of the material and the type of stress used to produce the strain. When the body is subject to a simple tension (or compression), the body being free to contract in a direction normal to the line of action of the stretching forces, k, i.e. the ratio $\dfrac{\text{stress}}{\text{strain}}$, is called ***Young's modulus.*** When the stress is due to shear the ratio $\dfrac{\text{stress}}{\text{strain}}$ is termed the ***modulus of rigidity*** of the material.

Referring to Fig. 6·3, if s is the area of the upper face of the block the shearing stress is $\dfrac{F}{s}$, and since the strain is $\tan\theta$, or θ (expressed in circular measure), if the angle of shear $\widehat{CBX}$ is small, the modulus of rigidity is $\dfrac{F}{s} \div \theta$. The above method of determining k for

shearing stresses is only applicable to india-rubber, for the angle of shear is usually so small that it cannot be measured directly. Other methods are therefore employed, but they are beyond the scope of the present work.

Experiment.—Obtain a rectangular block of indiarubber 20 cm. long, 4 cm. wide, and 5 cm. thick, and cement one of the 20 cm. × 4 cm. faces to a vertical wall, the long edge being vertical. Cement a thin metal plate to the face opposite that cemented to the wall and suspend various loads by means of a hook attached to the plate. Measure with the aid of a travelling microscope the descent of the plate for each load and calculate the mean descent for unit change in load. Calculate the rigidity, n, of indiarubber as indicated in the following example.

Example.—The mean extension for a change in load of 1 kgm. was 0·040 mm. for the above block. Find the modulus of rigidity n.

$$\therefore \text{Angle of shear} = \frac{0{\cdot}004}{5} \text{ radian.}$$

$$\text{Change in stress} = \frac{1000 \times 981}{20 \times 4} \text{ dyne. cm.}^{-2}$$

$$\therefore \text{Modulus of rigidity} = \left(\frac{1000 \times 981}{20 \times 4} \times \frac{5}{0{\cdot}004}\right) \text{ dyne. cm.}^{-2}$$

$$= 1{\cdot}53 \times 10^7 \text{ dyne. cm.}^{-2}$$

Young's Modulus.—Suppose that a wire of length L and radius r is stretched by a load of mass m, the wire being free to contract in a direction perpendicular to the stretching force. The stress in the wire is $\frac{mg}{\pi r^2}$, since the wire is subject to stretching forces equal in magnitude to the weight of the load. If l is the increase in length the strain is $\frac{l}{L}$, and the modulus of elasticity, denoted by Y in this instance, is given by

$$Y = \frac{\text{tensile stress}}{\text{tensile strain}} = \frac{mg}{\pi r^2} \div \frac{l}{L} = \frac{mgL}{\pi r^2 l}.$$

Experimental Determination of Young's Modulus.—Two identical wires of the material under investigation are suspended from a beam. The method of attaching the wires to this support is important, for if either wire slips the elongation will not be due to the load alone. One method is to pass one long piece of wire between two brass plates which are afterwards screwed together with the aid of two or more bolts and nuts—cf. Fig. 6·5. This wire is arranged so that the lengths of the two free portions are approximately equal. A scale graduated in mm. is screwed to the left-hand wire whilst a second wire carries a vernier and a scale-pan. In this way, since the wires are identical, any temperature change will affect each wire to the same extent, so that no differential

expansion due to temperature variations will be noticed. The wire carrying the mm. scale is slightly stretched by suspending from it a convenient load. The pan attached to the second wire is usually sufficient to keep it straight when it is otherwise unloaded. The initial reading of the scales having been noted, the wire carrying the vernier is suitably loaded and the scale and vernier reading observed. To determine the ratio $\frac{\text{stress}}{\text{strain}}$ it is advisable to increase the load by 500 gm. at a time and observe the scale reading after each increment has been made: a graph showing the relationship between the load [ordinate] and the elongation [abscissa] is then constructed. Since the elongation is proportional to the load if the elastic limit has not been exceeded, this graph should be a straight line. With a piece of black cotton as a guide, the best straight line should be drawn through the points on the diagram and the slope calculated [cf. p. 14]. If θ is the slope of this line, we have

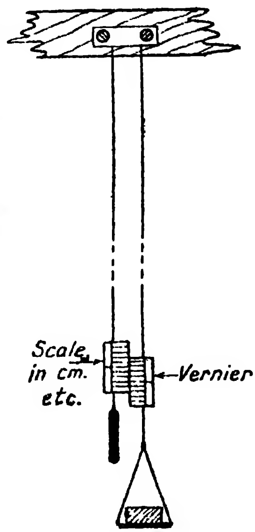


FIG. 6·5.—Apparatus for determining Y.M. for Metals in the form of Wires.

$$Y = \frac{mg}{\pi r^2} \cdot \frac{L}{l} = \frac{Lg}{\pi r^2} \cdot \theta.$$

Young's modulus can therefore be calculated if, in addition to the above observations, the length and mean radius of the wire are known. The mean radius is determined with the aid of a micrometer screw gauge [cf. p. 7]. To test whether or not the elastic limit has been exceeded observations should also be made as the load is removed; corresponding observations will be in agreement, the wire returning to its original length, if the elastic limit has not been passed and the mean value of the extension for each load should be used in constructing the above graph.

Searle's Apparatus for Determining Young's Modulus for the Material of a Long Wire.—Two wires of the same material are hung from the same rigid support, their lengths being about 2 metres. Each carries at its lower end a brass rectangular frame from the lower sides of which suitable loads may be supported. In Fig. 6·6, A and B are the wires while C and D represent an end-on view of these frames. E is one of two bars freely hinged to the frames so that one frame may be displaced relatively to the other. H is a metal strip, carrying a spirit level S, and freely moving about a fulcrum M at one end. At the other end it rests upon the point N of a vertical screw R, operated by the divided head T. The pitch of the screw is 0·5 mm. and the periphery of T is divided into 50 equal divisions. When the head T is rotated through one division its point moves 0·01 mm.

A load of 1 kgm. is applied to each wire so that they shall be straight and the reading of the screw observed when one end of the air bubble is at the centre of the level. [This permits the position of the bubble to be adjusted more precisely than if it is attempted to adjust the bubble to a central position.] The load on one wire is then increased by 1 kgm. so that the wire is stretched and the air bubble displaced. By rotating the screw this bubble may be brought back to its standard position. The amount by which the point of the screw is moved is equal to the extension of the wire. The load is then increased in stages up to a maximum, removed 1 kgm. at a time, and readings of the screw taken for each load. A graphical or other method is then used to determine the mean extension for an increase in load of 1 kgm. and a value for Young's modulus for the material of the wire calculated as in the previous experiment.

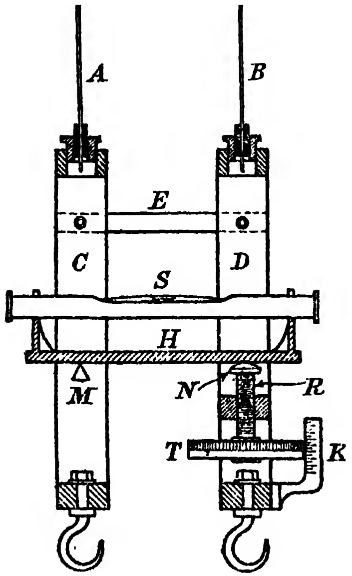


FIG. 6·6.—Searle's Apparatus for investigating the Stretching of Wires.

Poisson's Ratio.—When a wire is subjected to the action of stretching forces only, in addition to the elongation which occurs, there is a contraction in all directions perpendicular to the length of the wire. The change in diameter relative to the original diameter is termed the ***lateral contraction strain.*** The ratio of the lateral strain to the elongation strain is known as ***Poisson's ratio*** (σ). For indiarubber available in the form of a long solid tyre, this ratio may be determined from observations on the change in diameter and the change in length. For other substances Poisson's ratio is calculated from the other elastic constants for each particular substance. The necessary formulae are too difficult to prove here.

Volume Elasticity or Bulk Modulus.—We have seen that if a body is subjected to a uniform pressure its volume diminishes. If p is the increase in pressure necessary to cause a volume V of a material to diminish by an amount v, the stress is p, for a pressure is defined as a force per unit area [cf. p. 71], while the strain is $\frac{v}{V}$. The modulus of elasticity, which, in such an instance, is termed the ***volume elasticity*** or ***bulk modulus,*** is therefore $p \div \frac{v}{V}$, i.e. $\frac{pV}{v}$,

and is denoted[1] by β. The reciprocal of the bulk modulus is termed the ***compressibility*** of the substance, and is denoted by κ, so that $\kappa = \frac{1}{\beta}$.

The Compressibility of Liquids.—In an elementary account of hydrostatics it is always assumed that liquids are incompressible and, in fact, enormous pressures are required to alter by a small amount the volume of unit volume of all liquids, i.e. their compressibilities are always small. The fact that water was not a liquid with zero compressibility was first established by CANTON in 1762. BACON, at an earlier date, had subjected water, completely filling a lead sphere, to pressures greater than atmospheric, but the experiments were frustrated by the fact that the sphere always sprang a leak, or else the water escaped through the walls of the vessel which were porous. Canton used a glass vessel shaped like a thermometer and containing mercury, but with an open capillary tube. The level of the mercury in the stem of the instrument at a definite temperature was noted. The vessel was then heated until the mercury just filled it: the open end of the capillary was sealed and the instrument allowed to cool to its former temperature. It was found that the mercury stood at a higher level in the capillary than formerly. To account for this it might be assumed:—

(i) that the mercury had previously been compressed by the external air, or

(ii) that the vessel was reduced in size when the pressure inside was less than atmospheric.

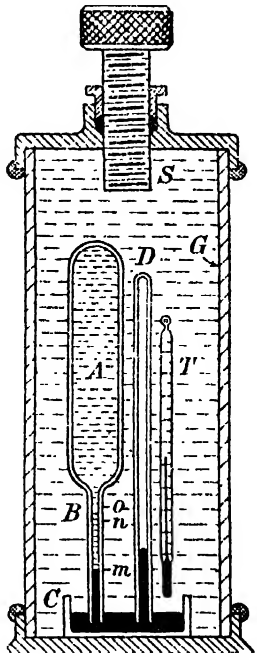


FIG. 6·7.—The Compressibility of Liquids [after Oersted].

The experiment was then repeated with water in the same vessel: the change in level of the water was greater than in the case of mercury. It was therefore established that water was a compressible substance.

Experimental determinations of the compressibilities of liquids are beset with many difficulties, but the underlying principles are shown by the following experiment due to OERSTED (1822). One form of his apparatus—an example of a class of instruments known as ***piezometers***—is shown in Fig. 6·7. The liquid under investi-

[1] In acoustics E is used instead of β.

gation was contained in a cylindrical glass vessel, A, provided with a capillary tube, B, dipping below the surface of mercury contained in a trough, C. D was a narrow uniform glass tube, closed at the top, whose lower end dipped into the same dish of mercury. It contained air. The whole was placed in a wide glass tube, G, completely filled with water. Packing glands prevented the escape of water from between the ends of G and the metal discs closing its ends. By rotating the screw H so that it moved downwards very large pressures were exerted inside G, and these were transmitted to the liquid in A. The change in pressure inside the apparatus was deduced from the change in volume of the air in D. ***Theory:*** Let V be the volume of A up to the zero mark 0 on its stem. Let v be the volume per division on the stem B. Suppose that when the pressure was p_1, the mercury in B stood at m: when it was p_2 at n. Then $(m-n)v$ is ***apparently*** the reduction in volume of a volume $(V+mv)$ of liquid when the pressure changes by an amount (p_2-p_1). The ***apparent compressibility*** is therefore given by

$$\kappa = \frac{\text{apparent diminution in volume}}{(\text{original volume}) \times (\text{change in pressure})} = \frac{(m-n)v}{(V+mv)(p_2-p_1)}.$$

The more exact theory, due to Lamé, shows that from the rise of mercury in B, together with other relevant data, the difference between the compressibilities of the liquid and glass may be deduced.

Regnault, in 1847, was the first person to obtain accurate values for the compressibilities of several common liquids, including mercury. In interpreting his observations he made use of the theoretical investigations of Lamé.

Experiment.—The smallness of the compressibility of water is shown by the following experiment. Water completely fills a metal box having no lid. When a small bullet is fired into one side of this box the volume of water is diminished by an amount practically equal to that of the bullet before the water has had time to rise. Consequently enormous pressures are exerted on the sides of the box and this bursts.

The Volume Elasticity or Bulk Modulus of an Ideal Gas at Constant Temperature.—Let P and V be the pressure and volume of a given mass of an ideal gas at constant temperature. Let the pressure become $P+p$, the corresponding volume being $V-v$. Then the increase in stress is p, while the corresponding strain is $\frac{v}{V}$, so that, by definition, the bulk modulus is $p \div \left(\frac{v}{V}\right)$. Since the gas is an ideal one and therefore obeys Boyle's law

$$(P+p)\,(V-v) = PV,$$

or

$$pV - vP - pv = 0.$$

If p and v are small compared with P and V their product may be neglected, so that

$$pV - vP = 0, \text{ or } P = \frac{pV}{v} = \beta.$$

The volume elasticity of an ideal gas at constant temperature is therefore equal to the pressure to which it is subjected. [N.B.—The pressure must be expressed in absolute or in gravitational units.]

Alternative Proof. Let P and V become P + δP and V + δV respectively. The increase in stress is δP, while the strain is $-\frac{\delta V}{V}$.

$$\therefore \text{ by definition, } \beta = \lim_{\delta P \to 0}\left[-\frac{\delta P}{\left(\frac{\delta V}{V}\right)}\right] = \lim_{\delta P \to 0}\left[-V\frac{\delta P}{\delta V}\right].$$

Now by Boyle's law, $(P + \delta P)(V + \delta V) = PV$.

$$\therefore V\delta P + P\delta V = 0,$$

since the product $\delta P . \delta V$ may be neglected. Hence

$$\lim_{\delta P \to 0}\left[-V . \frac{\delta P}{\delta V}\right] = P, \text{ i.e. } \beta = P.$$

Energy due to Strain.—In order to deform a body work must be done by the applied forces. The energy thus spent is stored in the body which is then said to possess ***strain energy***. This energy is lost when the stress is removed, appearing as heat, i.e. the body is temporarily at a temperature above that of its surroundings. The whole of the work done in deforming the body is only completely regained if its elastic limit has not been passed, for in this latter instance a permanent set is produced in the body and the energy necessary to do this is not regained when the stress is removed.

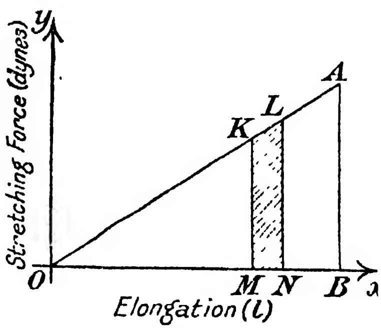


FIG. 6·8.—Energy due to Strain.

For a wire in which the stress does not exceed its elastic limit the amount of work done in stretching it may be calculated as follows :—If a point A, Fig. 6·8, represents the state of the wire when the stretching force is F and the elongation l, the work done in producing this condition is represented by the area of the triangle OAB. To prove this, consider two points K and L on OA, and draw KM and LN perpendicular to the x-axis. If the points K and L are very close together then during the deformation MN

the stretching force may be considered constant and equal to that represented by KM, so that the work done is represented by KM $\times$ MN, i.e. it is represented by the area of the rectangle KLNM. Similarly, every such small rectangle into which the triangle OAB may be divided represents a quantity of work done. The total work done in deforming the wire is represented by the sum of all these rectangles, i.e. the area OAB. The work represented by this area is $\frac{1}{2}F \times l$. If L is the length of the wire and r its radius, the strain energy per unit volume is

$$\tfrac{1}{2}Fl \div \pi r^2 L = \tfrac{1}{2} \cdot \frac{F}{\pi r^2} \cdot \frac{l}{L}.$$

This equation means that the strain energy per unit volume is one half the product of the stress and the strain.

The Behaviour of Solids when the Applied Stress exceeds their Elastic Limits.—(*a*) *Brittle Materials in Tension.* Cast iron, hardened iron, Portland cement, stone and brick are examples of a brittle substance. Fig. 6·9 (*a*) shows the relation between the strain and stress for such a substance. OA is linear, so that A is the elastic limit, but beyond A the graph is curved. The point B represents the stage when the substance breaks.

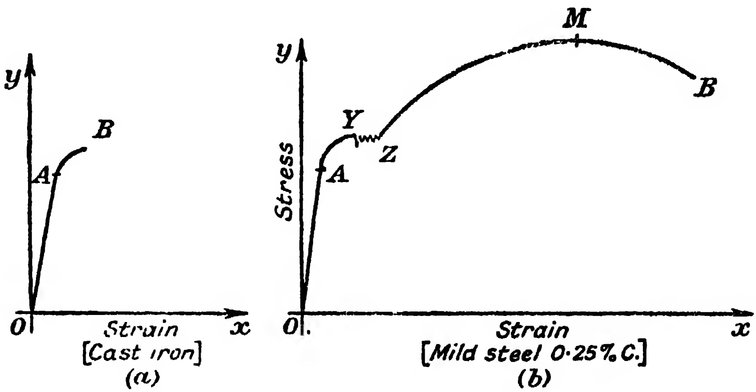


FIG. 6·9.—Behaviour of Solids when the Applied Stress exceeds their Elastic Limits.

(*b*) *Ductile Materials in Tension.* The stress-strain diagram for a 0·25 per cent. carbon steel, obtained autographically, is shown in Fig. 6·9 (*b*). A is the elastic limit and beyond this point the graph curves until the point Y is reached. Then comes the portion YZ of the curve, representing the stage during which there is a large increase in strain with practically no increase in stress : on a self-recording sensitive extensometer the portion YZ appears as an irregular wavy line, the stress corresponding to Z being less than

that at Y. Y is called the ***yield point,*** the corresponding stress being the ***yield-stress.*** In the stage AY the stretch is partly elastic and partly due to plastic flow in minute conglomerates of particles distributed throughout the material under test. Beyond Z the elongation becomes plastic: during the elastic stage the stretch is caused by simple tension but in the plastic stage strain becomes predominant so that the stretch is mainly due to total shear taking place throughout all parts of the specimen. As the stress is increased the stretch proceeds steadily until the bar is about to fracture. Then a stage with marked instability sets in and the piece becomes considerably thinner at one point, i.e. the specimen exhibits a local contraction and a marked roughening of the hitherto smooth machined surface of the material appears. The specimen exhibits a phenomenon known as 'necking.' Immediately this occurs the stress decreases automatically and the portion MB of the curve is obtained: the break finally occurs at B. The stress corresponding to M is called the ***ultimate strength*** or ***tensile strength*** of the material under test. Steel (0·2 per cent. carbon) has a tensile strength of 30 ton-wt. in.$^{-2}$, while among timbers, British Oak, 'that synonym for strength and durability,' with a tensile strength of about 7 ton-wt. in.$^{-2}$, stands supreme if certain foreign woods are excluded. Unfortunately, it contains acids which corrode iron and steel. It is for this reason that copper rivets are used in the construction of wooden ships.

Elastic Fatigue.—When a metal has been subjected to repeated alternations of stress it becomes 'fatigued,' i.e. its strength diminishes, which means that for a given stress the amount of strain increases. If the alternations are continued for a sufficiently long time the metal may ultimately develop a fracture.

In the manufacture of copper tubes of elliptical section the tubes are first drawn with a circular section. If the final operation of making the bore elliptical is carried out at once it is successful, but if the tube is allowed to remain overnight the process cannot be completed in the morning.

EXAMPLES VI

1.—Define the terms: *tensile stress, tensile strain, Young's modulus bulk modulus, compressibility.* Derive an expression for the bulk modulus of an ideal gas. Two uniform wires of the same material are such that the linear dimensions of one are double those of the other. If equal loads are suspended from the above wires calculate the ratio of the extensions produced.

2.—Derive an expression for the force inwards due to a rope under tension passing round a smooth curve. Calculate the limiting pressure inside a cylindrical boiler of 3 ft. radius, the sides being $\frac{1}{8}$ of an inch thick and made of a material which can stand a limiting pressure of 40 ton.-wt. in.$^{-2}$

3.—How would you proceed to determine Young's modulus for a substance in the form of a uniform wire? If Y.M. for steel is 2×10^{12} dyne. cm.$^{-2}$, what mass must be suspended from a steel wire 2 metres long and 1 mm. diameter to stretch it by 1 mm.?

4.—A solid has a volume of 3·5 litres when the external pressure is 1 atmosphere. If the bulk modulus of its material is 10^{12} dyne. cm.$^{-2}$, calculate the change in volume when the body is subjected to a pressure of 25 atmospheres.

5.—Explain Hooke's law and describe how you would proceed to verify it for the extension of a vertical wire under load. A copper wire, 2 metres long and 3 mm.^{2} cross-sectional area, is suspended vertically and a load of 5 kilograms attached to its lower end. Calculate the work done in stretching the wire if Young's modulus for copper is $1{\cdot}2 \times 10^{12}$ dyne. cm.$^{-2}$

6.—Define *Young's modulus* and the *modulus of bulk elasticity.* Calculate the value of the latter modulus for a substance of which 1 cubic decimetre is reduced in volume by 0·01 cm.^{3} by an increase of pressure of 20 atmospheres.

7.—Explain what is meant by the statement: 'Young's modulus for steel is 2×10^{12} dyne. cm.$^{-2}$' Calculate the mass of the load which must be suspended from a steel wire 1 mm. in diameter to produce an elongation equal to 0·2 per cent. of its original length.

8.—How would you compare experimentally the value of Young's modulus for copper with the value for the modulus of brass, being given wires of the same standard gauge?

9.—Calculate the modulus of bulk elasticity for a substance of which 1 cubic decimetre is reduced in volume by 0·004 cm.^{3} when subjected to an increase of pressure of 16 atmospheres.

10.—Calculate the density of water at the bottom of a lake 150 metres deep assuming that the compressibility of water is $\frac{1}{21{,}000}$ atmos.$^{-1}$

11.—Given that Young's modulus for steel is 2×10^{12} dyne. cm.$^{-2}$ calculate its value in pounds weight per square inch.

12.—A spiral spring of negligible mass is hung vertically and is such that a load of 6·5 gm.-wt. produces an extension of 10 cm. If the spring carrying a load of 508 gm.-wt. is pulled downward, show that the load will execute a S.H.M. when the spring is released, and determine its period.

13.—An elastic string of natural length $2a$ can just support a certain weight when it is stretched until its whole length is $3a$. One end of the string is now attached to a point in a smooth horizontal table, and the same weight is attached to the other end and can move on the table. Prove that if the weight is pulled out to any distance and then let go, the string will become slack again after a time $\frac{\pi}{2}\sqrt{\frac{a}{g}}$.—(L.I.)

14.—A mass of metal of volume 500 cm.^{3} hangs on the end of a wire whose upper end is rigidly fixed. The diameter of the wire is uniform and equal to 0·4 mm. and its Young's modulus 7×10^{11} dyne. cm.$^{-2}$ When the metal is completely immersed in water, the length of the wire is observed to change by 1 mm. Find the length of the wire if the acceleration due to gravity is 980 cm. sec.$^{-2}$—(N.H.S.C. '29).

PART II

HEAT

CHAPTER VII

THERMOMETRY

Temperature.—Our ideas of the terms 'hot' and 'cold' are based upon our sense of touch or of feeling. A hand placed near a fire experiences a different sensation from that arising from its immersion in snow. In this way different bodies can be arranged in such an order that as one passes from one body to the next the sensation experienced is one of greater cold. These degrees of heat and cold correspond to a certain state or condition of the object. The following experiment shows that our hand is not a reliable indicator of temperature.

Experiment.—Suppose that A, B, and C are three bowls containing cold, tepid, and hot water respectively. Place the left hand in A, and the right hand in C; after half a minute transfer both hands to B. It appears hot to the left hand but cold to the right.

Moreover, the human hand is not sufficiently sensitive to detect small changes in temperature, neither is it capable of withstanding extremes of temperature.

In order to fulfil these purposes, thermometers have been constructed. In these use is made of the change in some physical property of a substance which varies continuously with the temperature, e.g. the increase in the volume of a liquid, or of a gas at constant pressure, which generally takes place with rise in temperature. It is also necessary to define two temperatures so that a scale of temperature may be constructed. These two temperatures must be constant and easily reproducible at all times and places; or if they are not constant the manner in which they vary with external influences must be known. The first such temperature is that of melting ice [free from contaminations] which is defined as the zero of the centigrade scale of temperature. In order to produce any appreciable change in this temperature the ice must be subjected to a pressure of several atmospheres. Since the variations in atmo-

spheric pressure never amount to more than a few cm. of mercury, we may say that for most practical purposes the melting-point of ice is constant. The second temperature, defined as 100° C., is that of steam when the external pressure is ***one standard atmosphere.*** Now mercury expands when heated and the value of gravity varies over the surface of the earth, so that the pressure of 76 cm. of mercury must be recorded by a barometer at 0° C. in latitude 45° N. Since such conditions cannot easily be realized the reading of the barometer must be corrected for these variations. In addition, it is quite fortuitous if the pressure so corrected happens to be exactly 76 cm. of mercury when the thermometer is calibrated. But since,

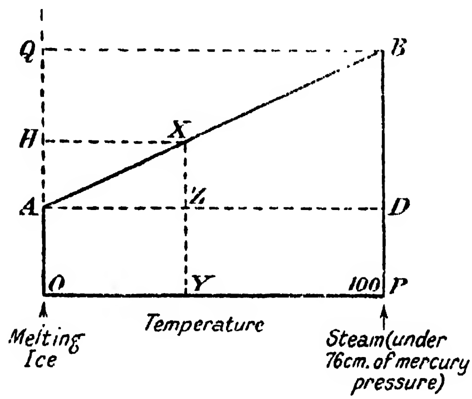


FIG. 7·1.—A Centigrade Scale of Temperature.

when the barometer reads about 76 cm., a change in pressure of 1 cm. of mercury causes a change of 0·37° C. in the boiling-point of water, the actual temperature associated with any particular pressure is easily calculated. Thus, if the corrected value of the observed pressure is 74·1 cm. of mercury, i.e. 1·9 cm. lower than the standard conditions, the boiling-point of water is

$$[100 - (0{\cdot}37 \times 1{\cdot}9)]^\circ \text{ C.} = 99{\cdot}30^\circ \text{ C.}$$

Here it must be emphasized that it is impossible, with any thermometer whatsoever, to measure the melting-point of ice or the standard steam temperature. The first is ***defined*** as 0° C., and the steam point under existing atmospheric conditions is found from observations on the barometer, the standard steam temperature being ***defined*** as 100° C. when the pressure is 76 cm. of mercury.

Scales of Temperature.—The early workers on thermometry used mercury-in-glass thermometers. The degree centigrade on such a thermometer is defined as follows :—When the temperature

of the thermometer *increases* by 1° C. the change in volume of the mercury is one-hundredth of the *decrease* in volume which occurs when the thermometer cools from the standard steam temperature to that of melting ice.

Thus suppose OP, Fig. 7·1, is a straight line 100 units long. Let OA and PB represent the lengths of the column of mercury in a mercury-in-glass thermometer, when the thermometer is at the temperature of melting ice and at the temperature of steam produced under standard conditions,[1] these lengths being measured from some arbitrary fiducial mark on the stem of the instrument. Through A draw a straight line parallel to OP to cut PB in D. Then DB could be divided in any arbitrary way to construct a scale of temperature. If it is divided into 100 equal divisions we shall have obtained a centigrade scale of temperature. To carry out this division conveniently we join AB by a straight line. Then if the length of the mercury column in the stem is OH, we obtain the temperature on the scale we have constructed by drawing HX parallel to OP to cut AB in X, and then drawing XZY normal to OP to cut AD in Z and OP in Y. Then

$$\frac{XZ}{OY} = \frac{BD}{OP},$$

i.e.

$$XZ = \frac{BD}{100}.OY.$$

Now $\frac{BD}{100}$ represents a change of 1°, so that the temperature is OY.

It should be very carefully noted that AB has been joined by a straight line simply for convenience, so that the statement that mercury is chosen as a thermometric substance because its expansion is uniform is a statement without meaning.

In England and English-speaking countries the Fahrenheit scale is employed for commercial and domestic purposes. On this scale the temperature of melting ice is 32° F., whilst the standard steam temperature is 212° F., i.e. the fundamental interval is divided into 180 equal parts. The reason for the adoption of these apparently arbitrary numbers is to be found in the fact that the zero on this scale was the lowest temperature that could be reached when the scale was proposed by Fahrenheit, viz. that of a mixture of snow and salt containing eutectic proportions [cf. p. 242]. The 100° F. was taken to be the temperature of a healthy person's body. It is now known that this temperature is 98·4° F.

[1] The thermometer must be placed in steam produced under pressures (i) just below, (ii) just above standard atmospheric pressure, and the position of the mercury under standard conditions found by interpolation. This remark applies to the standardization of all thermometers.

The other scale is due to Réaumur. On it the zero corresponds to the temperature of melting ice while the steam temperature is 80° R.

The Fundamental Interval.—The interval between the fixed points on a thermometer is called its fundamental interval. On the centigrade scale this is equal to 100 divisions; on the Fahrenheit 180; and on the Réaumur 80. Bearing in mind the magnitudes of these fundamental intervals, it is easy to convert the readings on one scale into those on the others.

Thus in order to convert a temperature of 35° C. into the corresponding temperatures on the Fahrenheit scale, 35 centigrade divisions equal $\frac{35 \times 180}{100} = 63$ Fahrenheit divisions; but the ice-point of the Fahrenheit scale is called 32, so that the required temperature is $(63 + 32)°$ F. $= 95°$ F.

Construction of a Mercury-in-Glass Thermometer.—A mercury-in-glass thermometer can be constructed from a long piece of uniform capillary tube, having sealed on at one end a piece of wider glass tubing and at the other a small funnel. A constriction is placed at A, Fig. 7·2, whilst just below there is a small bulb C. The whole is cleaned with aqua regia, alcohol, ether, acetic acid, more alcohol, and finally distilled water by passing these reagents in succession through the tube. It will be noticed that no solutions of solids, such as aqueous potassium bichromate, are used; this is because it is sometimes difficult to remove the solid particles which may have been present and become lodged in the capillary. The tube is finally dried by drawing air through it by means of a suction pump, the air having been passed over soda lime contained in a U-tube. Particles of soda lime are prevented from entering the tube by means of a swab of cotton-wool. The lower end is then warmed at B and finally closed, the end being rounded by gently blowing into the open end. To prevent the tubes from becoming contaminated during this procedure, a soda-lime tube may be attached to the open end of the thermometer. Mercury is then placed in the funnel above A, after which the tube and bulb are heated gently. The air is expelled in part, so that on cooling a little mercury enters the instrument. This mercury is boiled, all the air being expelled by the mercury vapour, so that when the source of heat is removed, as the vapour condenses, more mercury is drawn into the tube until it is completely filled. It is then maintained at a temperature a little above the highest to which the instrument will ever be

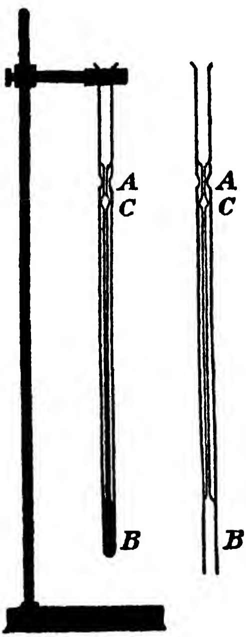


FIG. 7·2.—Construction of a Mercury-in-Glass Thermometer.

used and finally sealed off at A. The instrument is now allowed to cool slowly and left for a few days before being graduated—this is to allow the glass to contract to its original volume. In actual practice the glass goes on contracting for many years, and this helps to make the instrument difficult to use in precision work, because it implies an ever-changing zero.

Determination of the Steam and Ice Points on a Thermometer.—Of these two points on a mercury thermometer the upper one should always be found first; the ice point immediately afterwards. This is advisable because it is known that the glass continues to shrink after every heating so that it is better to take the ice point under definite conditions, viz. after every other reading of the thermometer.

The apparatus, shown in Fig. 7·3, is used for the determination of the steam point. It consists of a cylindrical vessel A, in which the water is boiled, surmounted by an open tube C. This tube is surrounded by a wider one having an outlet B for the steam. In precision determinations the U-tube D contains water to indicate any inequality between the pressure inside and outside A. The thermometer to be calibrated is supported by means of a cork and is so arranged that the final position of the mercury is just visible. When this position has become steady a small mark is made on the glass. The barometer is read and the temperature corresponding to this pressure is then calculated. [If vapour escapes from the tube E, more liquid must be placed in the boiler.]

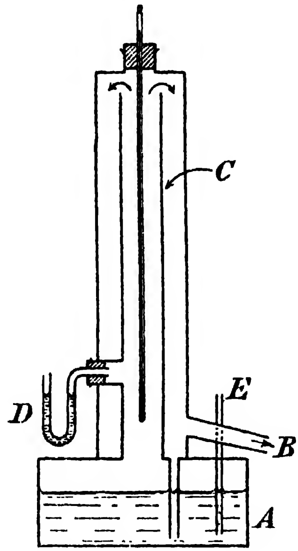


FIG. 7·3.—A Hypsometer.

The ice point is found by placing the thermometer in melting ice. The ice, or better, ice shavings, should be contained in a Dewar Flask and well stirred. In order to ensure that the ice is melting the whole should be covered with distilled water. It is more usual to place the thermometer in a funnel of melting ice and allow the water to drip away; this apparatus is condemned if any accuracy is desired, because the various pieces of ice are not in contact with the bulb, and, furthermore, the sharp points on the ice may exert a variable pressure on the thin-walled glass bulb and so violate steady conditions. These sharp points on the ice may also cause the bulb to fracture.

The Essentials of a Thermometric Substance.—The thermometric substance should have a large expansion for a small change in temperature; its indications must be consistent and easily observed; it should have a large working range; it should be easily procurable; it should not be easily contaminated; it should also rapidly assume the temperature to which it is subjected. Mercury possesses most of these characteristics, but mercury thermometers suffer from the following defects:—Mercury freezes at

$-40°$ C. so that for temperatures below $-30°$ C. an alcohol or a pentane thermometer must be used. Moreover, it is found that the mercury expands in jerks, a fact due to the change in volume of the bulb when subjected to changes in pressure caused by the variations in the angle of contact of the mercury. The density of mercury is high so that the volume of the bulb is affected by the changes in pressure produced by the variations in the length of the mercury column. In this respect alcohol and chloroform are to be preferred, but they cannot be used at high temperatures and there is a tendency for them to distil into the upper parts of the thermometer.

To increase the working range of mercury the space above the mercury may be filled with nitrogen. When the mercury expands the pressure inside increases so that the boiling-point of the mercury is raised. The thermometer may therefore be used to measure higher temperatures, but the variations in volume of the bulb due to pressure are increased so that such a thermometer cannot be regarded as a reliable instrument. It has also been proposed to use an alloy of sodium and potassium which is liquid from $-8°$ C. to 700° C., but it has been found that, as a result of the reduction of the glass by the alloy, a brown deposit of silicon is formed after a time so that the liquid cannot be seen.

The Errors of a Mercury Thermometer.—A thermometer is supposed to read 0° C. when placed in melting ice ; if it does not the reading must be observed and a correction applied. Similarly, the observed steam point will seldom be correct, so that a further correction is necessary. Suppose that the thermometer reads

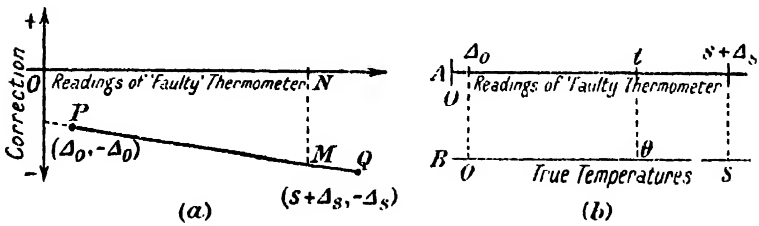


FIG. 7·4.—Graphical Method of Determining the Corrections to a Faulty Thermometer at Intermediate Temperatures.

Δ_0 and $(s + \Delta_s)$ when the temperature of its bulb is 0° C. and $s°$ C., respectively, s being the temperature of steam produced under existing conditions. Then the corrections are $-\Delta_0$ and $-\Delta_s$ respectively. Let these facts be represented graphically as follows, the aim being to express the true temperature as a function of the *readings* of the faulty thermometer. In Fig. 7·4 (*a*) the readings of the thermometer are taken as abscissæ while the corresponding corrections are the ordinates. Thus P and Q are the

points $(\Delta_0, -\Delta_0)$ and $(s + \Delta_s, -\Delta_s)$ respectively. Now consider Fig. 7·4 (*b*), in which the scales of the uncorrected thermometer, A, and a correct centigrade thermometer, B, are shown. Let t be the temperature as observed on A. Then in (*b*) the distance of t from the initial end of the line (scale) is $(t - \Delta_0)$. Let θ be the corresponding temperature on the scale B. Then

$$\frac{t - \Delta_0}{s + \Delta_s - \Delta_0} = \frac{\theta}{s},$$

since s° C. correspond to $(s + \Delta_s - \Delta_0)$ divisions on the actual thermometer.

This equation gives the true temperature θ as a function of the reading t. The correction y, or $\theta - t$, is given by

$$\theta - t = \frac{(t - \Delta_0)s}{s + \Delta_s - \Delta_0} - t,$$

i.e.

$$y = \frac{-s\Delta_0 + t(\Delta_0 - \Delta_s)}{s + \Delta_s - \Delta_0}.$$

But this is the equation to the straight line PQ in Fig. 7·3 (*a*), for this is

$$\frac{y + \Delta_0}{t + \Delta_0} = \frac{-\Delta_s + \Delta_0}{s + \Delta_s - \Delta_0}$$

which reduces to the above. Hence the ordinate NM at any point N on the t-axis in Fig. 7·4 (*a*) gives the correction at that point, i.e. at the observed temperature t.

In carrying out this construction due regard must be paid to the signs of Δ_0 and Δ_s.

This method is only justifiable if the bore of the thermometer is uniform. If it is not, and the thermometer is to be a standard one, it must be calibrated by breaking off a portion of the mercury column and observing its length [expressed in scale divisions] at various parts of the tube. The method of doing this is beyond the scope of this work, so that the simplest method of checking the indications of a thermometer is by comparing them with those of a standard instrument. Commercial mercury-in-glass thermometers are graduated by using subsidiary fixed points, and the divisions are generally not of equal size.

The most troublesome correction arises from the fact that all the mercury in the thermometer is not at the same temperature except in rare instances; it is therefore necessary to make a correction for stem exposure. A method of estimating this correction will be explained later [cf. p. 185].

Results obtained with mercury-in-glass thermometers are often vitiated by the fact that when the reading was being made, the line of sight was not normal to the mercury thread at its extremity,

i.e. errors due to parallax were not avoided. The use of a low power lens helps to minimize errors due to this cause.

JOULE made the first accurate mercury thermometer, and it was because of the pains he took in getting accurate readings with his thermometers that he made such remarkable discoveries in the science of thermodynamics.

The Paris Standard Thermometer.—The stem of this instrument is sometimes 1 metre long, although often it is only half this length. The bore of the tube is cylindrical so that the mercury shall move more regularly in the tube. [Cheap thermometers are often made with elliptical bores in order to facilitate seeing the mercury, but this is an objectionable practice in an instrument for scientific purposes for it increases the 'sticking effect' always associated with the motion of mercury over a glass surface.] The position of the mercury is observed with the aid of a microscope so that a change in temperature of 0·001° C. is detectable. Such an instrument is only suitable for measuring steady temperatures.

The Clinical Thermometer.—The clinical thermometer, Fig. 7·5 (*a*), is used for determining the temperature of the human body. It is a mercury-in-glass thermometer with a short working range, and is of the "maximum type," i.e. it registers the highest temperature to which its bulb has been exposed since re-setting.

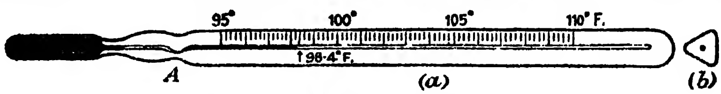


FIG. 7·5.—A Clinical Thermometer.

The range is 95° F. to 110° F, each degree division being divided into five equal parts. The earliest mercury thermometer used for the above purpose was not self-registering and the temperature was read while the thermometer was in the patient's mouth. In a later form there was a short column of mercury separated by an air bubble from the rest of the mercury. When the bulb of this thermometer was heated the short column of mercury moved forward, but it did not recede when the temperature fell. In the instrument in use to-day a small constriction, A, is placed in the bore of the thermometer near to the bulb—this is the special feature of this thermometer. The constriction must be such that mercury passes freely through it when the mercury in the bulb expands, but that it prevents the mercury in the stem above the constriction from returning when the temperature of the bulb falls; the indication of the thermometer may then be observed at leisure.

After use, the above instrument is re-set by shaking it so that the mercury in the stem is jerked downwards. Sometimes this operation is difficult so that frequently one finds clinical ther-

mometers with two constrictions at a short distance apart. These are not so small as in the usual instrument and, while sufficient to prevent the mercury returning to the bulb when the temperature of the latter falls, do not offer so great a resistance to the mercury when the latter is forced downwards by shaking the thermometer.

The temperature indicated by a clinical thermometer inserted in a patient's mouth, for example, does not reach a maximum value at once, for that part of the mouth in contact with the thermometer is cooled when the thermometer is inserted, and some time must elapse before the circulation of the blood restores the temperature to its original value at this point. To diminish this cooling effect and to make the thermometer quick in its response to temperature changes, it is essential to make the bulb small, and this, in turn, implies that the bore of the instrument must be small if an open scale of temperature is to be available. It is then difficult to locate the mercury in the stem unless it is provided with a 'lens front.' A section of the stem of such a thermometer is shown in Fig. 7·5 (*b*), and when the temperature recorded by the thermometer is being noted, the thermometer should be held in such a position that the mercury column is viewed directly through the sharp rounded edge of the stem. A magnified image of the thread is then seen.

The instruments are made in different sizes, known as 'half-minute,' 'one-minute,' etc. This time indicates the period after which the indications of the thermometer will have become steady when it is in use.

If at any time it becomes desirous to check the indications of a clinical thermometer, this is best done by means of a comparison with a standard instrument. Since clinical thermometers are not capable of showing a falling temperature, the comparison should be made by placing the two thermometers in a bath, the temperature of which is gradually rising, say 0·1° C. per minute. The bath is thoroughly stirred, and comparisons made at various points over the range of the instrument.

Maximum and Minimum Thermometers.—For meteorological purposes it is necessary to know the extremes of temperature reached over some period—generally a day. SIX's maximum and minimum thermometer, Fig. 7·6, is used for this purpose. A is a bulb filled with alcohol and connected to a second bulb partially filled with alcohol. The connecting tube is a capillary filled with mercury as shown. Two small steel indexes are placed above the mercury in E and F. These are supported by small springs. A fall in temperature causes the mercury to rise in E, so raising the index C whilst leaving D unaltered in position. A rise in temperature causes the

mercury in E to fall and in F to rise, so raising D, but leaving C in position. This is because the alcohol wets the indexes and passes by them, while the mercury does not. The lower ends of the indexes C and D give respectively the minimum and maximum temperatures to which the instrument has been subjected. The temperature scales are not exactly equal, since the mercury EF also expands with increase in temperature and affects the scale on the left-hand side.

When these thermometers are made the whole instrument is cooled down to a temperature beyond the lower limit of the scale, and the bulb on the left-hand limb sealed. The result is that when the instrument is in use there is a pressure above the liquid in B equal to the saturation pressure of the liquid at the appropriate temperature, plus the partial pressure of the enclosed air. It is very essential that some air should be contained in B so that the pressure at any point in the liquid in A, or the capillary tube E attached to it, should exceed the saturation vapour pressure of the liquid at the existing temperature. For suppose that the level of the mercury in F is below that in E—i.e. the instrument is at a low temperature. If a small bubble were formed in E it would grow unless the partial pressure of the air in B is greater than that due to a column of mercury equal in height to the difference between the mercury levels in E and F.

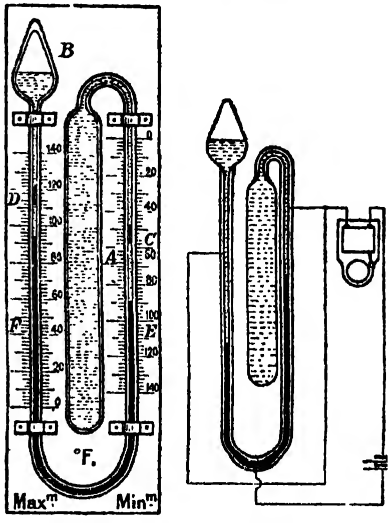


FIG. 7·6.—(a) Six's Maximum and Minimum Thermometer, (b) fitted with electrical device to give audible warning.

It has recently been found that the position of the zero of these instruments undergoes considerable changes if the alcohol contains acetone.

Fig. 7·6 (b) shows a thermometer of this type fitted with three platinum contacts as shown connected to a battery and bell. If the temperature passes beyond the limits appropriate to the position of the contacts, audible warning is given.

The Beckmann Thermometer.—This particular type of thermometer, which is very frequently used in physical chemistry experiments to determine the molecular weights of dissolved substances, was designed to measure small differences of temperature accurately. In order to make an ordinary thermometer sufficiently sensitive for this purpose its bulb would have to be large and its stem very long. This latter condition is very undesirable since such stems are easily fractured.

Now in a Beckmann thermometer Fig. 7·7, the bulb, A, is large but the stem is only of the usual length. This thermometer will not record actual temperatures but only differences of temperature. The stem is divided into six divisions each corresponding to 1° C., and each is further subdivided into one hundred parts. In order to render the instrument useful over a considerable range of temperature a small reservoir at B contains mercury which can be added to that already in the bulb, or more mercury can be abstracted from the bulb and left in this reservoir. Suppose, for example, one wishes to measure small changes of temperature in the region near to 30° C.; the thermometer is first inverted so that the mercury forms one continuous column. Its bulb is then placed in a water bath at a few degrees above this temperature and the thermometer is gently tapped. The mercury column breaks at a point near to the top of the small reservoir, so that when the temperature falls to 30° C. the mercury level should be on the scale of the thermometer. If this condition has not been obtained the above process must be repeated. A change in temperature may then be measured in the usual way.

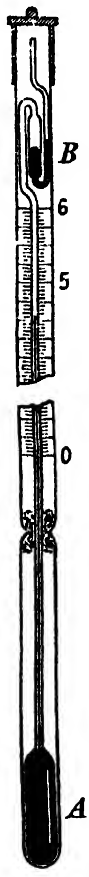


Fig. 7·7.—A Beckmann Thermometer.

Strain Thermometers.—These thermometers which are only of historical interest depend upon the fact that a heterogeneous body changes its shape considerably when subjected to changes in temperature. In one form three strips of platinum, gold, and silver, respectively, are made into a single ribbon by passing them through a rolling mill. The gold is placed between the two other metals. The whole is coiled into a spiral with the most expansible metal [Pt] inside. The spiral unwinds itself when its temperature is raised, the amount of twist being measured by a pointer attached to one end of the spiral, the other end being fixed.

Hot Surface Thermometers.—To measure the temperature of a surface with any degree of accuracy is not easy, and yet it often happens that such measurements are necessary. To measure the temperature of a stationary or slowly moving surface the bulb of the thermometer is coated with copper and attached to a flat piece of copper which has been gold-plated so that it shall be a poor radiator of heat. This is placed in contact with the surface whose temperature is required and, since the copper retains any heat imparted to it, the temperature indicated is the temperature of the surface. For fast-moving surfaces, such as calender rollers in paper-making, etc., the copper shoe is arranged so that it just avoids contact in order to prevent friction, the distance being approximately $\frac{1}{100}$th of an inch from the surface of the roller. To effect this clearance the shoe carries a cross-piece which is the common axis for two wheels. The outer edges of these wheels are in contact with the roller and the diameters of the wheels are sufficient to effect the necessary clearance.

EXAMPLES VII

1.—Calculate the boiling-point of water when the barometric height is 74·9 cm. of mercury. If a thermometer reads 99·2° C. under these conditions, what is the correction to be applied to it ?

2.—Describe a clinical thermometer, and the procedure you would adopt in order to check its indications.

3.—Describe a thermometer suitable for measuring the maximum and minimum temperatures of a greenhouse. How would you test its accuracy ?

CHAPTER VIII

THE EXPANSION OF SOLIDS

The Expansion of Solids.—When a body is heated it usually expands, i.e. its volume increases. Substances such as fused silica and invar steel expand by only a very small amount when they are heated, whereas gases expand much more rapidly when heated under constant pressure. The effect of heating a material is strikingly shown by hanging a piece of nickel wire, about 0·5 mm. in diameter, from a hook and stretching it vertically by means of a small weight. An electric current of about 10 amperes is passed through the wire so that it becomes red hot. The rapid descent of the weight, and its return when the current is broken, is a vivid manifestation that such a body expands when heated.

Let us consider the following: suppose it is required to use the expansion of a given rod to construct a centigrade scale of temperature. It is unlikely that the scale so constructed will be identical with any other centigrade scale of temperature and therefore we shall denote temperatures on this scale by ϕ. To obtain the scale we measure l_0, the length of the rod at the temperature of melting ice and l_{100}, its length when it is at the temperature of steam produced under a pressure of one standard atmosphere. To do this we may have to measure its length when it is in steam produced on two different occasions preferably when the pressure is (a) below, (b) above that of a standard atmosphere. Then $(l_{100} - l_0)$ is the expansion of the rod caused by a rise in temperature of 100 degrees and when the expansion is $\frac{1}{100}(l_{100} - l_0)$ we have a rise in temperature of one degree. Let l_ϕ be the length of the rod when its temperature is ϕ, then

$$\frac{l_\phi - l_0}{l_{100} - l_0} = \frac{\phi}{100},$$

or

$$\frac{l_\phi}{l_0} - 1 = \phi\left(\frac{\frac{l_{100}}{l_0} - 1}{100}\right).$$

If we put $\dfrac{\phi}{100}\left(\dfrac{l_{100}}{l_0} - 1\right) = \kappa$, we have

$$l_\phi = l_0(1 + \kappa\phi),$$

where κ is known as ***the coefficient of linear expansion of the material of the rod.***

If the length of the rod when its temperature is t on some other centigrade scale (e.g. that of mercury-in-glass) is l_t, then it is found that

$$l_t = l_0(1 + \lambda t + \mu\lambda^2),$$

and higher terms in t may be added in very accurate work. In practice it is found that μ is very small and under these circumstances the above equation becomes

$$l_t = l_0(1 + \lambda t),$$

i.e.

$$\kappa = \lambda,$$

and we often write

$$l = l_0(1 + \lambda t),$$

where l is the length of the rod at temperature t. Then

$$\lambda = \frac{l - l_0}{l_0 t}.$$

Since $(l - l_0)/l_0$ is a number, the dimensions of λ are those of the reciprocal of a temperature. This is an important point, for it shows that λ depends on the scale of temperature adopted in any experiment portending to determine λ. Thus, if λ_C and λ_F are the coefficients of linear expansion for a given material when centigrade and Fahrenheit scales of temperature are used, then

$$\lambda_C = \frac{l - l_0}{100 l_0} \text{ and } \lambda_F = \frac{l - l_0}{180 l_0},$$

where l and l_0 are the lengths of the rod at the temperatures of steam produced under standard conditions and of melting ice respectively. Hence

$$\lambda_F = \tfrac{5}{9}\lambda_C.$$

In exactly the same way the volume expansion of a substance between the temperatures of melting ice and that of steam produced under a pressure of one standard atmosphere may be used to construct a centigrade scale of temperature when we should have

$$v_\phi = v_0(1 + \eta\phi),$$

where η is ***the coefficient of cubical expansion of the substance.***

Using some other centigrade scale of temperature we should find

$$v_t = v_0(1 + \alpha t + \beta t^2),$$

and in practice $\beta \rightarrow 0$, so that $\alpha = \eta$. We then have

$$v = v_0(1 + \alpha t).$$

[The coefficient of linear expansion at a given temperature t of a substance in the form of a rod whose length is l at the temperature t, is

$$\frac{1}{l} \cdot \frac{dl}{dt},$$

and this must be calculated when the relation between l and t has been found experimentally.

Similarly, the coefficient of volume expansion is

$$\frac{1}{v} \cdot \frac{dv}{dt}.]$$

If a unit cube, i.e. a cube whose edge is 1 cm., of an isotropic solid material is heated 1° C. each edge becomes $(1 + \lambda)$ in length, where λ is the coefficient of linear expansion. Under these conditions the original volume of 1 cm.^{3} will have become $(1 + \lambda)^3$ cm.^{3} and this is equal to $1 + 3\lambda + 3\lambda^2 + \lambda^3$. Now, since λ is small, λ^2 and λ^3 will be much smaller; it is therefore justifiable to neglect them, so that the final volume $= 1 + 3\lambda$. From this it is seen that the increase in volume is 3λ for a rise in temperature of 1° C., and this has been styled the coefficient of volume expansion: hence the coefficient of volume expansion of an isotropic solid material is equal to three times the coefficient of linear expansion.

Determination of the Coefficient of Linear Expansion of a Metal in the Form of a Tube.—The apparatus consists of a brass tube, AB, Fig. 8·1, about 1 metre long and 0·5 cm. diameter. A brass collar about 0·5 cm. long is soldered near each end of the tube. A small steel ball-bearing is soldered to the collar near B, whilst a needle-point is similarly attached

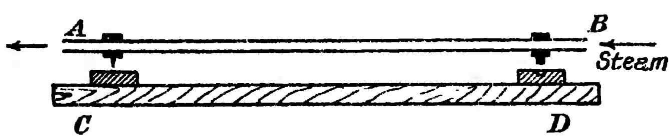


FIG. 8·1.

to the collar near A. Two brass plates are then screwed to a wooden board CD. The plate at D has a small cavity to receive the sphere while the needle-point rests on a second plate near C. A stream of cold water is passed through the tube, the temperature of which is observed with the aid of a calibrated thermometer. While the sphere rests in its socket a scratch is made by the pointer on the piece of brass below it. The water supply is turned off, the thermometer removed, and a copious supply of steam passed

through the tube. During this procedure the brass tube AB is placed at a distance of several feet from the board so that the distance between the two plates on CD does not alter. The temperature of the tube having been ascertained from observations of the barometric pressure, the tube is supported by dusters and the sphere B placed in the socket provided for it. A second scratch is then made on the plate at C. This plate is removed and the distance between the two scratches measured with a vernier microscope. Let Δl be this distance; let l_1 be the length of the tube at its initial temperature t_1. If the steam temperature is t_2 we have

$$l_1 = l_0 (1 + \lambda t_1) \text{ and } l_1 + \Delta l = l_2 = l_0 (1 + \lambda t_2)$$

Whence
$$\frac{l_2}{l_1} = \frac{1 + \lambda t_2}{1 + \lambda t_1}$$

so that λ may be calculated. λ is sometimes called the zero coefficient of linear expansion to distinguish it from $\bar{\lambda}$ the mean coefficient of linear expansion between temperatures t_1 and t_2, which is defined by the equation

$$\bar{\lambda} = \Big[\lambda\Big]_{t_1}^{t_2} = \frac{\Delta l}{l_1(t_2 - t_1)}. \quad \text{[cf. p. 171].}$$

The student should convince himself of the reality of the difference between these two coefficients by performing such an experiment and making the appropriate calculations.

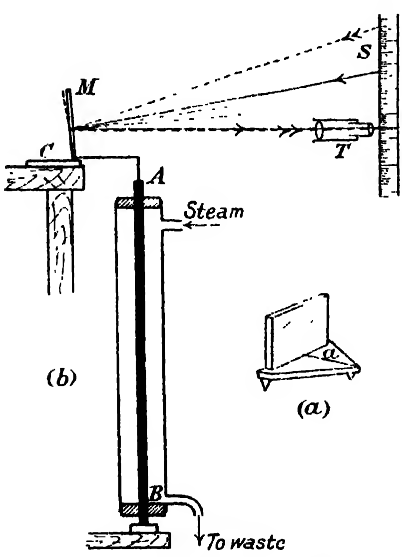


FIG. 8·2.—Optical Lever Method of measuring the Expansion of a Rod of Metal.

The Optical Lever: Determination of the Coefficient of Linear Expansion of the Material of a Rod.—The optical lever, as here used, consists of a small triangular piece of brass provided with three short legs at its corners, and having a plane mirror at right-angles to its base—Fig. 8·2 (*a*). To use the lever to measure the expansion of a rod due to a change in temperature, the latter is mounted vertically in a glass tube through which steam may be passed. The lower end of the rod rests on a brass plate, while one leg of the optical lever rests in a small indentation on the top of the rod—cf. Fig. 8·2 (*b*). The other legs rest on a brass plate, C. T is a telescope, and S a vertical scale in cm., etc., these being arranged on

a common normal to the mirror so that an image of the scale is sharply focussed on the crosswires of the telescope.

The particular division on S whose image is on the cross-wires is noted when the rod has been left for some time at room temperature, t_1, as observed by means of a thermometer. Steam is then passed through the tube surrounding the rod (the temperature, t_2, being deduced from the barometer reading); the rod expands and the mirror is tilted. When steady conditions have been obtained, the scale reading seen on the cross-wires is noted. If Δl is the actual expansion of the rod, the angle of tilt of the lever is $\Delta l/a$, where a is the distance indicated. Suppose that d is the difference of the readings as observed by T. If D is the distance of the scale from the mirror, then

$$\frac{d}{D} = 2\frac{\Delta l}{a},$$

since if a mirror rotates through an angle θ, a ray of light incident upon it rotates through 2θ [cf. p. 347].

If the length of the rod is measured, the zero coefficient of expansion for the material of the rod may be deduced from the equation

$$\frac{l_2}{l_1} = \frac{l_1 + \Delta l}{l_1} = \frac{1 + \lambda t_2}{1 + \lambda t_1},$$

where the symbols have their usual meanings.

To check the value so obtained, and see that the apparatus has not been disturbed, it should be allowed to cool to room temperature and the scale reading seen in T compared with that obtained originally.

The Comparator Method.—This method was designed by the International Committee of Weights and Measures at Paris for the purpose of comparing the length of any metre scale at various temperatures, with that of a standard metre maintained at constant temperature. Two massive stone pillars carry vertical microscopes, Fig. 8·3, each fitted with a micrometer eye-piece, the distance between the microscopes being approximately one metre. The standard metre is placed in one trough and the scale under examination placed parallel to the standard in a second trough. To assist in maintaining the bars at constant temperature the troughs are double-walled, the bars being placed in the inner compartments, and water from thermostats circulates between the walls of these troughs. The temperatures of the baths are given by carefully calibrated thermometers, efficient stirrers being employed to maintain a uniform temperature in each trough. The two troughs rest on wheels so that they may be moved along rails supported on a mass of concrete. In this way, first the standard metre, and then the other, is brought under the microscopes.

When the standard metre is below the microscopes these are displaced laterally so that the images of the fiducial marks on the bar coincide with the cross-wires in the eye-pieces of the micrometers. These will be called the 'zero positions' of the micrometers.

The distance between the cross-wires is then 1 metre provided that the axes of the microscopes are vertical and the surface of the bar is horizontal. The experimental bar is then placed in the above position and the shift given to each micrometer to establish coincidence between the image of each mark on the bar and the cross-wires of the microscope through which it is observed recorded. Let these shifts be α_0 cm. and β_0 cm. respectively: these quantities are considered positive if the microscopes have to be moved outwards. The distance between the marks on the bar is therefore

$$(100 + \alpha_0 + \beta_0) \text{ cm.} = l_0 \text{ cm. (say).}$$

The temperature of this bar is then altered to $t°$ by varying the temperature of the water flowing round the trough in which it is

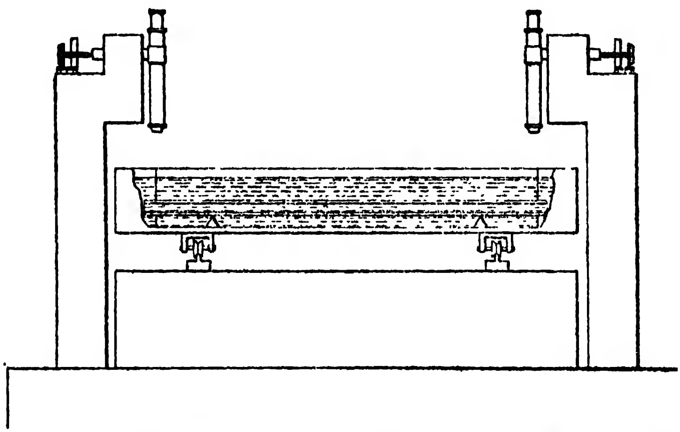


FIG. 8·3.—Comparator Method for Measuring the Linear Expansion of Rods.

situated. Let the shifts given to the micrometers measured from their zero positions to re-establish 'coincidence' be α_t cm. and β_t cm. respectively. Then $l_t = (100 + \alpha_t + \beta_t)$ and the coefficient of linear expansion for the material of the bar may be deduced for $l_t = l_0(1+\lambda t)$, or

$$\lambda = \frac{(\alpha_t - \alpha_0) + (\beta_t - \beta_0)}{(100 + \alpha_0 + \beta_0).t}.$$

Before doing this, however, the standard metre should again be brought below the microscopes to see whether the positions of the pillars bearing the microscopes have varied; if they have a correction must be applied.

Example.—A certain distance measured by a scale in cm., etc., is H cm. The temperature of the scale is $t_2°$ C. If the scale had been divided correctly at $t_1°$ C., what is the true value of the distance?

At t_1° C. each division is 1 cm. long,

$$\text{i.e. } 1 = l_0(1 + \lambda t_1)$$

where λ is the coefficient of linear expansion for the material of the rod and l_0 is the distance between two consecutive cm. marks at 0° C.

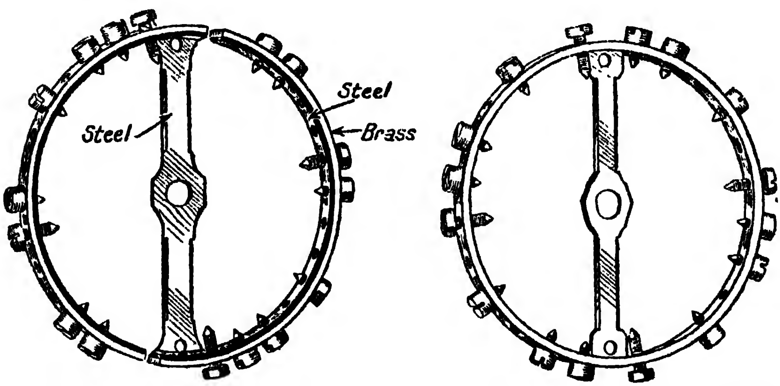
At t_2° C. the distance between two consecutive cm. marks is

$$l_0(1 + \lambda t_2) = \frac{1}{1 + \lambda t_1} \cdot (1 + \lambda t_2) \simeq [1 + \lambda(t_2 - t_1)]$$

Hence the distance required is $H[1 + \lambda(t_2 - t_1)]$ cm.

Some Consequences of Expansion.—Industry often makes use of the expansion of metals. The iron tyres of cart-wheels are fitted while they are red hot; when cold, their grip is considerably increased. The barrel of a gun consists of coaxial cylinders, the outer ones of which are in turn shrunk on to the remainder. Greater resistance to shock is thus obtained.

The rate of working of chronometers and watches is controlled by the oscillation of a balance wheel under the influence of a 'hair spring.' An increase in the diameter of the wheel causes



(*a*) Bimetallic Balance Wheel. (*b*) Monometallic Balance Wheel (not cut).

FIG 8·4.

it to oscillate more slowly, while an increase in stiffness of the spring makes the wheel oscillate more quickly. Now a rise in temperature increases the diameter of the wheel but reduces the rigidity of the spring, both of which tend to augment the periodic time of the wheel, i.e. the chronometer 'loses.' To compensate for this the rim of the wheel is constructed in at least two parts brass and steel frequently being used—Fig. 8·4 (*a*). The more expansible metal [brass] is placed on the outside so that a rise in temperature causes the section to curl inwards whereby the effective diameter is reduced.

Recent work by GOULD, in America, has shown that the above type of balance wheel may be more than effectively replaced by those made from 'elinvar', a nickel-steel alloy, whose coefficient of linear expansion is small—Fig. 8·4 (*b*). Moreover, the hair spring should also be constructed from this same material, since its elastic properties are practically unaffected by changes in temperature. Other advantages of using elinvar will be mentioned later.

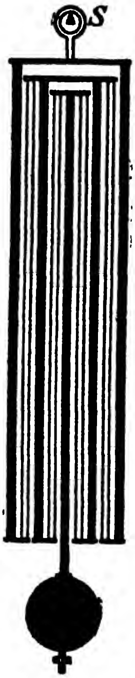


FIG. 8·5. — Harrison's Compensated Pendulum.

Pendulums of invariable length, and therefore constant periodic time, were first constructed by HARRISON. The manner in which this was accomplished will be gathered from Fig. 8·5. Five rods of steel were used in conjunction with four brass ones. When the whole is suspended from a fixed support any expansion of the steel rods increases the length of the pendulum, while that of the brass reduces it. For the compensation to be complete the expansion due to *three* steel rods, plus that of the short piece from S to the cross-piece, must be equal to that of *two* brass ones owing to the particular arrangement adopted.

Invar Steel.—About thirty years ago M. GUILLAUME discovered an alloy of steel and nickel [36 per cent. Ni] whose coefficient of expansion is very small. This particular alloy is known as *Invar*. It has been used in the construction of invariable pendulums, and for surveyors' tapes. These tapes may be calibrated at the National Physical Laboratory where they are immersed in a long trough through which water at a known temperature passes. A comparator method is used for measuring the length of the tape. When invar was first discovered it was thought that it would be a suitable material from which to construct standards of length. Recent work at the National Physical Laboratory, however, has revealed the fact that invar continues to 'grow' for many years after it has been manufactured; it is therefore not suitable for this purpose.

An artificially aged [1] metre bar has been kept under observation at the National Physical Laboratory for thirty years. In that period it has increased by 0·025 mm., and is still increasing at the rate of about 0·00025 mm. per annum.

Some years ago a new alloy described as 'stable' invar was introduced, but a four-metre bar of this material has been under observation since 1925 and in nine years has increased by 0·022 mm. Recently, the National Research Council of Canada has reported that a one-metre scale of an alloy known as 'Fixinvar' has contracted by 0·0009 mm. in nine months.

[1] The early growth of invar steel may be accelerated by a process of artificial ageing and this improves the subsequent stability.

EXAMPLES VIII

1.—A glass rod is 2·1605 metre. long at 0° C. and 2·1624 at 117 ° C. What is its mean coefficient of linear expansion between these temperatures ?

2.—Deduce the relationship between the coefficients of linear and volume expansion with temperature. Explain how to determine the coefficient of expansion of a liquid by weighing a solid of known expansibility in it.

3.—How would you proceed in order to test the accuracy of the statement—'the coefficient of linear expansion of brass is 0·000020 deg.$^{-1}$ C.' ? A simple pendulum consists of a bob suspended by a fine brass wire. The pendulum makes 3,600 vibrations per hour when the temperature is 15° C. Calculate what the period of the pendulum would be if the temperature fell to − 5° C.

4.—The relation between the volume and temperature of a substance is expressed by the equation

$$V_t = V_0[1 + 0{\cdot}000172\ t + 0{\cdot}0000021\ t^2].$$

Calculate the mean coefficient of expansion between 0° C. and 100° C. and the coefficient of expansion at 50° C.

CHAPTER IX

THE EXPANSION OF LIQUIDS AND GASES

The Expansion of Liquids.—Since a liquid has no definite shape of its own, but assumes that of the vessel in which it is contained, we cannot speak of its linear expansion but only of its volume expansion. When a liquid in a graduated container expands, the expansion observed is the expansion of the liquid together with that of the containing vessel—it is termed the ***apparent*** expansion as distinct from the real expansion of the liquid itself.

Experiment.—A flask and vertical tube leading from it are filled with coloured water so that the liquid stands about half-way up the tube. The flask is plunged into boiling water, when it is found that the level of the liquid in the tube falls temporarily, after which it rises. The fall is due to the sudden expansion of the glass before the liquid has had time to become heated. This experiment proves quite definitely that the expansion of a liquid is influenced by that of the container.

The Coefficient of Apparent Expansion of a Liquid.—Imagine that A, Fig. 9·1, is a vessel completely filled at $t_1°$ with liquid ; B represents the same vessel at $t_2°$. The vessel is now a little larger, but it is filled with liquid. C represents the state of affairs when the liquid

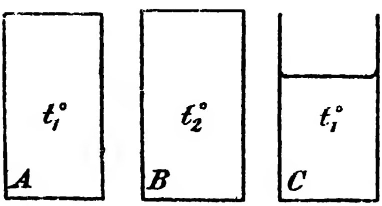


FIG. 9·1.

remaining in B and the vessel have cooled to $t_1°$. Let m_1 be the mass of liquid in A at $t_1°$, while m_2 is the mass left in B, i.e. the mass filling the vessel at $t_2°$. Then a mass $m_1 - m_2$ has been expelled. A brief glance at the diagram shows that it is a *mass* of liquid m_2 at temperature $t_1°$ which has expanded and driven out a *mass* $m_1 - m_2$ when the temperature was raised. The mass driven out is proportional to the change in volume of a volume represented

by a mass m_2. Since the change in temperature was $t_2 - t_1$ it follows that the mean coefficient of apparent expansion between these temperatures is

$$\frac{m_1 - m_2}{m_2(t_2 - t_1)}.$$

The Absolute Coefficient of Expansion of a Liquid.—This coefficient is determined from observations on the density of the liquid at different temperatures. If ρ is the density, and v the specific volume, i.e. the volume of one gram of substance,

$$\rho = \frac{\text{mass}}{\text{volume}} = \frac{1}{v}.$$

Hence, when the temperature is 0° C., $\rho_0 = \frac{1}{v_0}$; similarly, at t° C., $\rho_t = \frac{1}{v_t}$. But $v_t = v_0(1 + at)$, where a is the mean absolute coefficient of cubical expansion over a range in temperature from 0° C. to t° C. Hence

$$\frac{\rho_0}{\rho_t} = \frac{1}{v_0} \cdot v_0(1 + at) = (1 + at).$$

Hence, if the density of the liquid is measured at each of two temperatures, a may be deduced.

Note on the Mean Coefficient of Expansion.—The coefficient of volume expansion at a temperature t is defined by the equation

$$\alpha = \frac{1}{v}\frac{dv}{dt}.$$

Hence $\bar{\alpha}$, the mean coefficient of expansion between temperatures t_1 and t_2, is given by

$$\bar{\alpha} = \frac{\int_{t_1}^{t_2} \alpha\, dt}{\int_{t_1}^{t_2} dt} = \frac{\int_{v_1}^{v_2} \frac{dv}{v}}{t_2 - t_1} = \frac{\log_e \frac{v_2}{v_1}}{t_2 - t_1}.$$

This equation is exact, but if we write $\frac{v_2}{v_1} = 1 + \frac{v_2 - v_1}{v_1}$ and expand the logarithm as a series we have

$$\log_e \frac{v_2}{v_1} = \frac{v_2 - v_1}{v_1} + \frac{1}{2}\left(\frac{v_2 - v_1}{v_1}\right)^2 + \ldots$$

Hence, neglecting terms in $\left(\frac{v_2 - v_1}{v_1}\right)^2$ and higher, we have

$$\bar{\alpha} = \frac{v_2 - v_1}{v_1(t_2 - t_1)} = \left(\frac{\rho_1}{\rho_2} - 1\right)\frac{1}{(t_2 - t_1)},$$

where ρ, as usual, denotes density.

Now a better approximation is obtained by using the series

$$\log_e \frac{1+x}{1-x} = 2\left[x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots\right]$$

when $x = \frac{(v_2 - v_1)}{(v_1 + v_2)}$. Hence neglecting only terms in x^3 and higher, we have

$$\bar{\alpha} = \frac{v_2 - v_1}{\frac{1}{2}(v_1 + v_2)(t_2 - t_1)}$$

$$= \left(\frac{\rho_1 - \rho_2}{\rho_1 + \rho_2}\right) \cdot \frac{2}{(t_2 - t_1)}.$$

Indirect Determination of the Absolute Coefficient of Expansion of Liquids

The Weight Thermometer or Dilatometer.—The weight (or better *mass*) dilatometer is generally used for determining the absolute coefficient of expansion of a liquid indirectly—the method is an indirect one since a correction involving a knowledge of the coefficient of volume expansion of the material of the envelope or vessel (usually glass) containing the liquid has to be applied. The weight thermometer consists of a cylindrical bulb drawn out at one end into a fine capillary (but with thick walls), this latter being bent twice at right angles. The mass of the instrument is first found and it is then filled with liquid. This is done by heating the bulb gently on all sides to expel some air, allowing it to cool with its open end under the liquid, when some liquid is drawn into the bulb, and then proceeding as in the experiment when a mercury thermometer was constructed. When filled, the instrument is placed in melting ice, its neck still being immersed in a small reservoir containing the liquid. The dilatometer is removed after ten minutes, when we may assume that its temperature is that of melting ice, and its mass determined after its surface has been dried. To prevent a loss of liquid during the weighing operations a small glass receptacle of known mass may be attached below the neck of the instrument and the whole weighed together. The dilatometer is then placed in a beaker containing water or some other liquid at a constant known temperature; after ten minutes immersion its mass is determined when it is cold and its exterior dry.

Let V be the volume of the vessel and γ the mean coefficient of volume expansion of its material over the range of temperature employed. Let m be the mass of the liquid in the dilatometer and let t be the increase in temperature. Let ρ be the density of the liquid. Suffixes denote corresponding values of these variables at different temperatures.

Then

$$V_0\rho_0 = m_0.$$
$$V_t\rho_t = m_t.$$

But
$$V_t = V_0(1 + \gamma t), \text{ and } \rho_t = \frac{\rho_0}{1 + at}.$$

Hence
$$\frac{m_0}{m_t} = \frac{V_0\rho_0(1 + at)}{V_0\rho_0(1 + \gamma t)} = \frac{1 + at}{1 + \gamma t},$$

or
$$a = \frac{m_0 - m_t}{m_t t} + \frac{m_0}{m_t}\gamma.$$

Since γ is very small and m_0 is approximately the same as m_t, the above equation may be written

$$a = \frac{m_0 - m_t}{m_t t} + \gamma.$$

We have seen that the fraction $\frac{m_0 - m_t}{m_t t}$, which represents the mass *expelled* divided by the product of the *mass left in* and the rise in temperature, is the apparent or relative coefficient of expansion of the liquid in glass; hence we have proved that the absolute coefficient of expansion = the apparent coefficient of expansion + the coefficient of volume expansion of the material of the containing vessel.

The above method is, as here described, not a precision method since the temperature of the exposed stem is not the same as that of the beaker and it is difficult to estimate the necessary correction. Moreover, we have to assume that the expansion of glass is the same in all directions when calculating the volume expansion of glass from the linear coefficient. Actually, glass is a very anisotropic substance, i.e. its properties are not the same in all directions. However, the method can be made a precision one, but the details do not concern us here.

The Volume Dilatometer.—This method of determining the absolute expansion of a liquid has one advantage over that just described, viz. the correction for stem exposure is zero since the whole of the instrument can be raised to one and the same temperature. In addition, although the dilatometer may be filled by alternately heating and cooling, it may also be filled by a method in which the liquid is not heated, an expedient which is very desirable when dealing with inflammable liquids or a liquid which decomposes on heating to high temperatures. The dilatometer consists of a bulb, B, Fig. 9·2, having a graduated capillary CD attached to it; at D this opens out into a wider tube to receive any liquid if occasion arises. To fill the instrument it is fitted through a cork as indicated in Fig. 9·2. The liquid to be introduced into the bulb is contained in a wide tube, F, projecting

from the side of E into which the cork is inserted. A second tube, A, passing through the cork allows the apparatus to be connected to a vacuum pump so that it may be exhausted. The tap in A is then closed and the apparatus inverted. When air is slowly admitted the liquid is forced into the bulb B.

Before the expansion of the liquid can be found it is necessary to know the volumes of the various portions of the dilatometer. Let M_m be the mass of mercury required to fill the instrument to the m-th division on the stem when all is at 0° C. Let M_n have a similar meaning. Then the volume from one scale division to the next is

$$(M_m - M_n) \div (m - n)\rho_0,$$

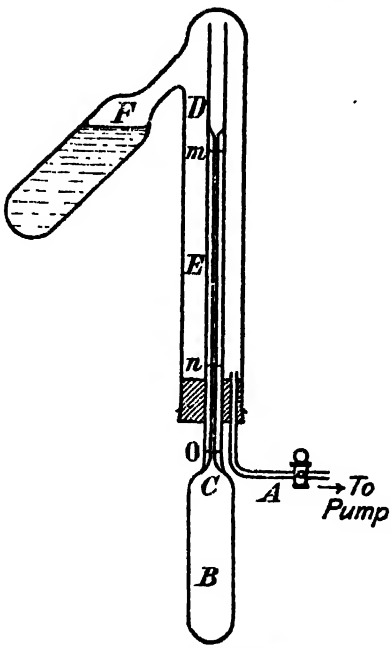


FIG. 9·2.—A Volume Dilatometer.

if the stem is of uniform bore and ρ_0 is the density of mercury at 0_0 C. Let $x = \frac{(M_m - M_n)}{\rho_0}$ Then the volume of the dilatometer up to the zero mark at 0° C. is

$$\frac{M_n}{\rho_0} - \left(\frac{n}{m - n}\right)x.$$

Let us assume that when the instrument contains the liquid under investigation that it is filled to the p-th division at 0° C. Then the volume of the liquid at this temperature is

$$\frac{M_n}{\rho_0} + \left(\frac{p - n}{m - n}\right)x.$$

At t° C., when the liquid extends to the q-th division, the volume of the dilatometer to this mark and therefore of the liquid at this temperature is

$$\left[\frac{M_n}{\rho_0} + \left(\frac{q - n}{m - n}\right)x\right][1 + \gamma t]$$

where γ is the coefficient of volume expansion of glass. The absolute coefficient of expansion of the liquid, α, is therefore expressed by

$$\left[\frac{M_n}{\rho_0} + \left(\frac{q - n}{m - n}\right)x\right][1 + \gamma t] = \left[\frac{M_n}{\rho_0} + \left(\frac{p - n}{m - n}\right)x\right][1 + \alpha t]$$

since, in general,

$$V_t = V_0(1 + \alpha t).$$

Expansion of a Liquid by Hydrostatic Methods.—(*a*) One method depends upon measurements of the apparent loss in mass when a body is suspended in a liquid. In order to increase the ratio of this apparent loss in mass to the actual mass of the body or sinker, it should be large and have a small mean density.

Such a sinker is indicated in Fig. 9·3. It consists of a hermetically sealed glass bulb weighted with mercury or lead shot so that it just sinks in the liquid under investigation. A hook is provided with which to suspend the sinker from the pan of a balance. Let M be the mass of the sinker in air, m_1 the mass in water at initial temperature t_1. [The liquid may conveniently be placed in a wide-mouthed Dewar flask and well stirred before observations are made. Changes in temperature will then be very small.] Then

$$M - m_1 = \text{mass of liquid displaced} = V_1\rho_1,$$

where V_1 is the volume of the bulb and therefore of the liquid displaced, and ρ_1 is the density of the liquid at t_1.

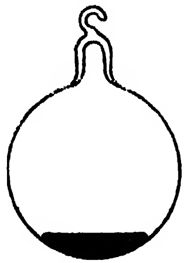


Fig. 9·3.

Similarly,

$$M - m_2 = V_2\rho_2 = V_1[1 + \gamma(t_2 - t_1)]\,\rho_2.$$

where γ is the coefficient of volume expansion of glass. [Strictly speaking γ is not the coefficient of volume expansion of glass defined in terms of an initial temperature 0° C., but for ordinary work the correction on this account is unimportant.]

Let $\bar{\alpha}$ the mean coefficient of expansion of the liquid over the range $t_2 - t_1$.

$$\text{Then } \bar{\alpha} = \frac{v_2 - v_1}{v_1(t_2 - t_1)} = \frac{1}{(t_2 - t_1)}\left(\frac{\rho_1}{\rho_2} - 1\right)$$

$$= \frac{1}{(t_2 - t_1)}\left[\frac{M - m_1}{M - m_2}\{1 + \gamma(t_2 - t_1)\} - 1\right].$$

A more accurate value of α, as defined on p. 172, is given by

$$\bar{\alpha} = \frac{\dfrac{M - m_1}{1 + \gamma t_1} - \dfrac{(M - m_2)}{1 + \gamma t_2}}{\dfrac{M - m_1}{1 + \gamma t_1} + \dfrac{M - m_2}{1 + \gamma t_2}} \cdot \frac{2}{(t_2 - t_1)},$$

and if $\gamma t \rightarrow 0$, the above expression may be simplified in the usual way.

(*b*) A second method consists in floating at two temperatures a hydrometer of known mass and expansibility in the liquid. If V_1 and V_2 are the volumes of the instrument at temperatures t_1 and t_2 respectively, β the coefficient of volume expansion of its material, ρ_1 and ρ_2 the densities of the liquid at these same temperatures, m_1 and m_2 the masses required to sink the hydrometer to the same fiducial mark in each instance, then

$$M + m_1 = V_1\rho_1,$$

and

$$M + m_2 = V_2\rho_2 = V_1[1 + \beta(t_2 - t_1)]\rho_2.$$

Then, $\bar{\alpha}$, the mean coefficient of expansion over the range of temperature t_1 to t_2, is given by

$$\bar{\alpha} = \frac{v_2 - v_1}{v_1(t_2 - t_1)} = \left(\frac{\rho_1}{\rho_2} - 1\right)\frac{1}{(t_2 - t_1)}$$

$$= \frac{1}{(t_2 - t_1)}\left[\frac{M + m_1}{M + m_2}\{1 + \overline{\gamma(t_2 - t_1)}\} - 1\right].$$

Direct Determination of the Coefficient of Absolute Expansion of a Liquid

The Method of Balancing Columns.—The methods hitherto discussed for determining the coefficient of expansion of a liquid all suffer from the defect that they involve a knowledge of the expansion of the material of the containing vessel. Dulong and Petit, about 1816, first developed a method for determining the coefficient of expansion of a liquid directly, i.e. a knowledge of the expansion of the material of the vessel was not involved. The principle is to compare directly the densities of the liquid at two temperatures between which the mean coefficient of expansion is required. The liquid was placed in a U-tube, ABCD, Fig. 9·4, one limb of which was kept cool (in melting ice) while the other was maintained at the desired temperature (steam).[1] Each limb was about 55 cm. long.

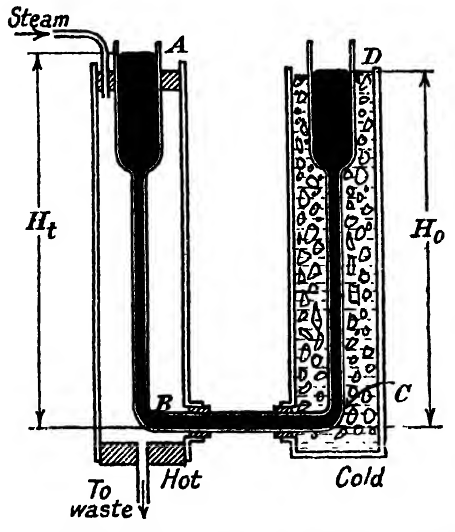


Fig. 9·4.—Principle of Dulong and Petit's Apparatus for Determining Directly the Expansion of Mercury.

Let t_0 and t be the temperatures of the cold and hot limbs respect-

[1] Actually, by using heated liquids in the vessel surrounding AB, a series of results was obtained, the temperature of the hot bath being given by an air thermometer.

ively. Then, when equilibrium has been reached, a column of liquid at t_0 balances a column of the same liquid at t. If Π is the atmospheric pressure[1], that at B is $\Pi + g\rho_t H_t$. At C it is $\Pi + g\rho_0 H_0$. Since these are equal

$$g\rho_0 H_0 = g\rho_t H_t.$$

$$\therefore \frac{H_t}{H_0} = \frac{\rho_0}{\rho_t} = 1 + \alpha t,$$

or

$$\alpha = \frac{H_t - H_0}{H_0 t}.$$

Hence, if H_t and H_0 are determined experimentally (by means of a cathetometer), a value for α may be deduced.

Strictly speaking, in this experiment, equilibrium is never established, for, on account of density differences in the liquid, there will always be two feeble currents in the cross-tube—an upper one from the hot limb to the cold one, and a lower one in the opposite direction. At the level of the axis of the tube a state of equilibrium may be considered to exist, and it is for this reason that the heights H_0 and H_t were measured from the axis of the cross-tube. To reduce the effects just referred to, the cross-tube is made narrow.

To emphasize the fundamental principles of this method of measuring the coefficient of expansion of a liquid the design of the above apparatus has been kept as simple as possible. For example, it has been assumed that the hot limb was enclosed in a vapour bath. Actually, it was placed in a copper vessel containing oil; this was heated by a furnace. Moreover, the temperature of the hot limb was measured by an air thermometer and by a weight thermometer. Consistent results were only obtained with the former, so that the indications of the weight thermometer were discarded.

This simple form of Dulong and Petit's apparatus is open to the criticisms that the temperature of the liquid in either limb was not constant, and that $(H_t - H_0)$ was not determined directly. It is desirable to do this since the accuracy of the final result depends chiefly on the accuracy with which $(H_t - H_0)$ is determined. The apparatus was improved by Dulong and Petit themselves, also by REGNAULT, and by CALLENDAR and MOSS. The work of these last three investigators will now be described.

Regnault's Apparatus.—The apparatus shown in Fig. 9·5 is a schematic representation of Regnault's. The tubes AB and CD,

[1] Expressed in the same units as $g\rho H$.

each 1·5 metres long, were connected by a horizontal tube AC, the connecting tube BD being bent to form a U-tube. The whole circuit was filled with mercury except for a small region in the U-tube, which was connected to an air pump, thus enabling the pressure in it to be increased until mercury was just about to escape from the hole K. CD was surrounded by melting ice, whilst AB was immersed in an oil bath, the temperature of which was taken by means of an air thermometer P, the bulb of which extended almost from the top to the bottom of the bath. From observations on the heights H_t, H_0, h_0, and h_0', the coefficient of absolute expansion was calculated. If $t°$ C. is the temperature of the hot column and 0° C. the temperature at every other point, the pressure at M, due to the mercury in AB, and the atmospheric pressure, Π, is $\Pi + g\rho_t H_t - g\rho_0 h_0'$, where g is the acceleration due to gravity, and ρ_t and ρ_0 are the densities of the mercury at the two temperatures. Similarly, the pressure at N is $\Pi + g\rho_0(H_0 - h_0)$, and this must be equal to the pressure of the air in MGHN, and hence to the pressure at M

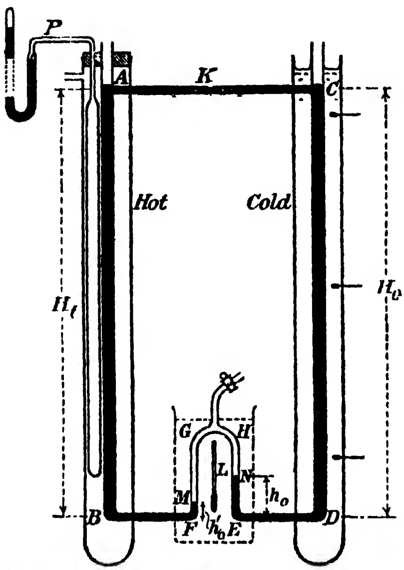


FIG. 9·5.—Regnault's Apparatus for Determining the Coefficient of Expansion of Mercury.

[N.B. No thermometers are needed if the 'cold' baths contain melting ice and the 'hot' bath is at the temperature of steam produced at atmospheric pressure.]

$$\therefore \rho_0(H_0 - h_0) = \rho_t H_t - \rho_0 h_0',$$

or

$$\rho_0(H_0 - h_0 + h_0') = \rho_t H_t.$$

$$\therefore \frac{\rho_0}{\rho_t} = \frac{H_t}{H_0 - (h_0 - h_0')} = 1 + at.$$

Now H_0, H_t, and $(h_0 - h_0') = \Delta h$ (say), are the three heights actually measured.

Hence

$$\alpha = \frac{H_t - H_0 + \Delta h}{(H_0 - \Delta h)t}.$$

Callendar and Moss' Apparatus.[1]—Two vertical tubes,[2] AB, A′B′, Fig. 9·6 (*a*), each nearly two metres long, were bent twice at right angles so that the portions BC, B′C′, were horizontal. The tube AA′ was made narrow to diminish the circulation of mercury from one vertical tube to the other. A mechanically driven paddle forced water, cooled to 0° C. by ice round M, through the wide tube surrounding AB. A′B′ was surrounded by an oil bath heated by an electric current passing through the loop of wire Q, which was made in the form indicated to distribute the heat energy in the bath. A second paddle R caused this oil to circulate steadily round A′B′. The temperatures of the baths were indicated by platinum thermometers P and P′ the bulbs of which extended the whole length of the baths, so that the mean temperature of each bath was known accurately. The tubes CD and C′D′ were also at 0° C. Special precautions were taken to keep the portions of the tubes strictly horizontal where they emerged from the baths. To prevent the conduction of heat along the horizontal tubes each array of tubes was silver-soldered to a brass block through which ice-cold water passed—cf. Fig. 9·6 (*b*). The heights of the longer columns were measured with steel tapes, carefully calibrated, while the difference, D′D, was measured with the aid

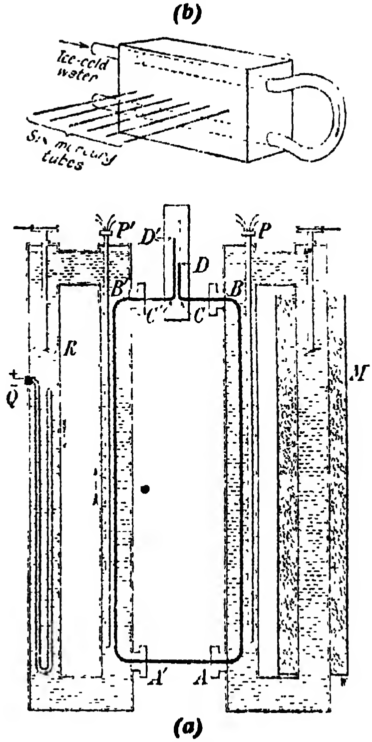


FIG. 9·6.—Apparatus for Investigating the Expansion of Mercury (Callendar and Moss).

[1] The essential features only are described. Moreover it is assumed that the temperature is 0° C. at all points except in the hot bath. Actually this was not so, but the corrections are too involved to be considered here.

[2] Actually there were six pairs of hot and cold columns placed in series. Successive columns were alternately hot and cold. The difference of level measured, DD′, was then six times that due to a single pair of hot and cold columns. This difference amounted to about 20 cm.

of a cathetometer. This consists of a horizontal telescope, having cross-wires in the eye-piece, and moving up and down a vertical graduated bar.

Let H_t and H_0 be the lengths of A′B′ and AB at temperatures $t°$ and 0° respectively. Let h_0 and h_0' be the lengths CD and C′D′ when both these columns are at 0°. If ρ_t and ρ_0 are the densities of mercury at $t°$ and 0° respectively, the pressures at A and A′ are $\Pi + g\rho_0 h_0 + g\rho_0 H_0$ and $\Pi + g\rho_0 h_0' + g\rho_t H_t$, where Π is the atmospheric pressure. Hence

$$\rho_0(H_0 + h_0) = \rho_t H_t + \rho_0 h_0'$$

$$= \left(\frac{\rho_0}{1 + at}\right) . H_t + \rho_0 h_0'.$$

$$\therefore\ a = \frac{H_t - H_0 + (h_0' - h_0)}{[H_0 + (h_0 - h_0')]t}.$$

If we call $h_0' - h_0$, the difference in levels DD′, which was actually measured, Δh,

$$\alpha = \frac{H_t - H_0 + \Delta h}{(H_0 - \Delta h)t}.$$

Callendar found that the mean value of a between 0° C. and 100° C. was $1{\cdot}82 \times 10^{-4}$ deg.$^{-1}$ C., and that a increased as the temperature increased.

The Anomalous Expansion of Water.—Water has a maximum density at about 4° C., a fact which shows that the expansion of water with rise of temperature is anomalous, i.e. water at 4° C. expands when it cools. To determine the temperature at which water has a maximum density Hope (1805) devised and carried out an experiment on the following lines. A, Fig. 9·7 (*a*), is a metal vessel, narrower at the central region than elsewhere. The central portion may be surrounded by a mixture of ice and salt at a temperature of about − 6° C. The upper section of the apparatus is coated with a thick layer of paraffin wax, B, while the lower portion is fitted with a Dewar flask, C. These are necessary to diminish the exchange of heat between the water, which is placed in the apparatus, and the external surroundings, except where the water is being cooled by the mixture in the trough, D. E is a thick piece of glass provided with two apertures through which pass two mercury thermometers F and K. This latter thermometer is constructed so that its zero mark is just outside the apparatus; it has a working range of about 20° C.

Initially the apparatus is filled with water at about 10° C. Before the annular trough is filled with the cooling mixture the reading of the upper thermometer will be slightly in excess of that of the lower one, for the warmer and therefore less dense portions of

the water are on top. When the mixture is applied the temperature of the water in the lower parts of the apparatus will begin to fall, at first slowly, but then more rapidly, and finally more slowly, until it is 4° C. Meanwhile, the water in the upper parts of the apparatus is cooled by the process of conduction, for heat flows downwards from this water to that surrounded by the freezing mixture in D, and warmed by heat received from the surroundings. The upper thermometer indicates the resultant effect. The cooling in the lower parts has been brought about by convection.

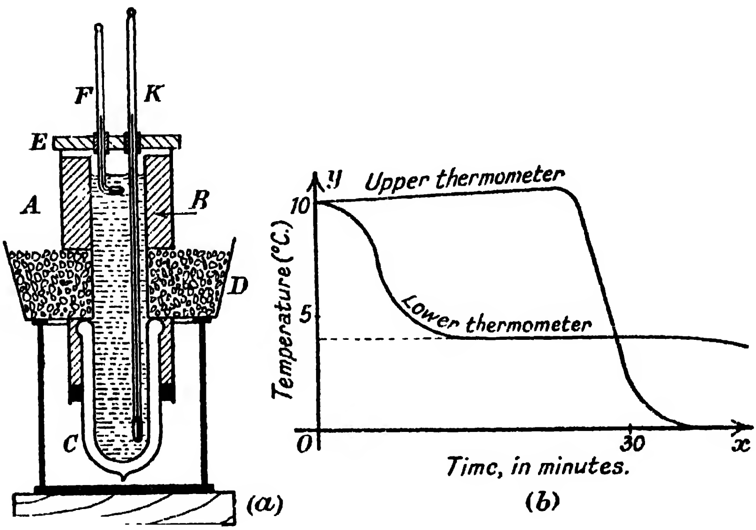
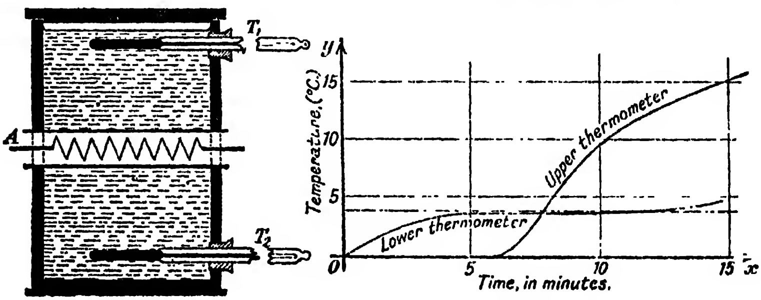


FIG. 9·7.—Hope's Apparatus. Modern Form.

After this stage has been reached the water in the central region becomes cooled to 0° C., but it ***does not rise,*** since water at a temperature between 0° C. and 4° C. has a density greater than that at about 10° C., which is still practically the temperature in the upper portions of the apparatus. More heat is then abstracted from the water near the centre, ice crystals are formed, and these rise. The water in the upper parts is cooled by the crystals as they melt until the temperature is reduced to 0° C. More ice crystals are then formed and these collect at the top, forming a layer of ice. The temperature indicated by the lower thermometer remains 4° C., although it tends to fall—due to heat lost by conduction.

If the temperatures of the two portions of water are plotted against time, curves similar to those shown in Fig. 9·7 (*b*) are obtained. This experiment proves that water has a maximum density in the neighbourhood of 4° C.

Hope's Experiment Modified.—DYSON has recently described the following experiment to demonstrate the fact that water has a maximum density at 4° C. The modification reverses the usual procedure, and works by warming ice-cold water by means of the energy dissipated in a small electric heater, A, Fig. 9·8 (*a*) fixed near the middle of a small vessel. This consists of a rectangular vessel 15 cm. × 7 cm. × 2·6 cm. Uniform heating of the surroundings is prevented by a poorly conducting covering to the apparatus. The walls are made of ebonite sheet about 6 mm. thick, the joints being made water-tight with the aid of Chatterton's compound. The heating coil is of nichrome tape from an old electric iron. It is wound on a narrow strip of mica, and protected at the back and front by wider mica strips. The resistance of this coil is about 10 ohms. The coil is mounted in a thin-walled copper tube. The procedure is to fill the apparatus with ice-cold water. This is left for a minute or two, then removed, and the whole refilled with ice-cold water. After eddy currents have subsided, the current is switched on (0·75 ampere) and readings of two mercury-in-glass thermometers, T_1 and T_2, situated as indicated, are noted at half-



(*a*) (*b*)

FIG. 9·8.—Modified Form of Hope's Apparatus.

minute intervals. The water in the central portion of the apparatus becomes warmed and until its temperature is greater than 4° C. sinks, displacing the water in the lower part. This occurs because water between 0° C. and 4° C. has a density greater than that of water at 0° C. As the heating proceeds, however, the water in the central region finally attains a temperature of 4° C., but it does not rise, since the density of the water in the superincumbent layers is less than unity. No convection currents are produced in the upper part of the apparatus until the temperature of the water near to the heating coil exceeds about 8° C., for then the density of the water close to the heating coil is less than that above it, and so convection currents are formed. These tend to increase as the heating proceeds. In this argument the effect of the heat exchange between the apparatus and its surroundings has been neglected—in practice this exchange will modify slightly the shape of the ideal curves shown in Fig. 9·8 (*b*). The advantages of this apparatus are that it is quick in action and no freezing mixture is required.

If this experiment were continued for some time the temperature of the lower thermometer would rise above 4° C., owing to heat being conducted downwards, but there would then be no convection currents in this part of the apparatus since the water at the top is always hotter.

Further Experiments on the Maximum Density of Water.—JOULE and PLAYFAIR (1851) investigated the temperature at which the density of water is a maximum in the following way, and the result they obtained is more reliable than that obtained with Hope's apparatus. Two vessels, A and B, Fig. 9·9, made of tinned iron and filled with air-free distilled water, were connected at the bottom by a brass pipe, C, and accurately ground stop-cock, D, whilst at the top they were joined by a rectangular trough, E, 6 in. long and 1 in. deep. A slide placed in this trough when necessary prevents the flow of water from one vessel to the other. The cylinders themselves were each 6 in. in diameter and 4 ft. 6 in. long. They were supported in two places by means of wooden brackets, H_1, H_2, and hay-bands wrapped round the vessels prevented the exchange of heat between the vessels and the surroundings from being excessive. To keep the apparatus free from vibration, it was allowed to rest on a support not in contact with the floor of the laboratory.

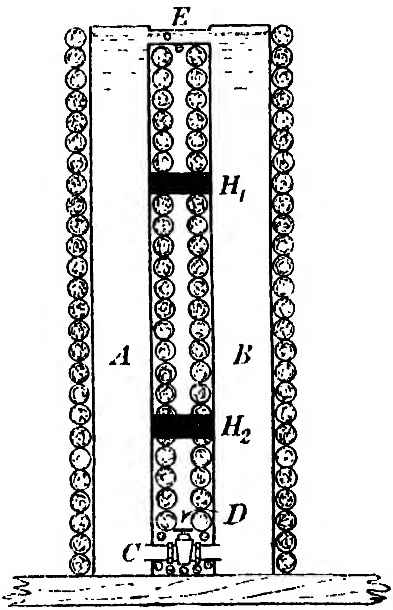


FIG. 9·9.—Joule and Playfair's Apparatus for investigating the Temperature when the Density of Water is a Maximum.

When the stop-cock D was opened and the slide carefully removed, a flow of water took place from one vessel to the other if there was the least difference between the density of the water in the two cylinders. This flow was made manifest by placing a hollow glass bead or ball in the iron trough. The mass of this bead was such that it only just floated—'a matter of great importance, as the slightest buoyancy is accompanied by a certain degree of capillary attraction, and makes the ball liable to adhere to the sides of the trough.' The temperatures were determined with the aid of mercury thermometers, sufficiently sensitive to detect changes in temperature of less than 0·005° C.

In making an experiment with this apparatus, the stop-cock in the connecting tube was closed, the water in each vessel thoroughly stirred; when it had come to rest, the stop-cock was opened, the slide removed, and the motion of the bead observed. If this

moved, it indicated that the water in the cylinder towards which the bead moved had the greater density. When a pair of different temperatures had been found for which the density of the water was the same, then one of them must be above and the other below the temperature at which water has a maximum density. Joule and Playfair obtained a series of such pairs of temperatures in which

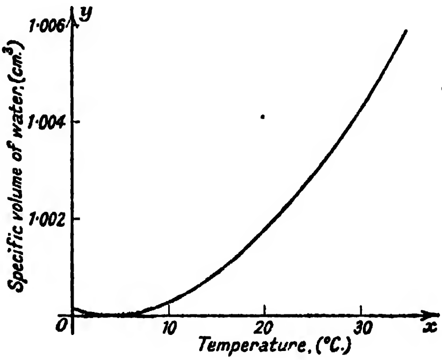


FIG. 9·10.—The Specific Volume of Water. Its Variation with Temperature.

the temperature difference became smaller and smaller. In this way they located the maximum density of water at 3·95° C.

The variations in the specific volume of water with change in temperature have been investigated at the *Physikalisch-Technischen Reichsanstalt*, Berlin, where the method of 'balancing columns' was used. The results, confined to the range 0° C. — 40° C., are shown in Fig. 9·10.

The Millilitre.—Originally the kilogram was defined as the mass of a cubic decimetre of pure water at the temperature of its maximum density (and under a pressure of one standard atmosphere). The underlying idea was that there should be a simple relation between the unit of mass and the unit of volume. Having agreed to this definition several French physicists were entrusted with the work of constructing a standard kilogram of platinum. Before the middle of the last century it had been definitely established that the mass of the above standard was not identical with that of a cubic decimetre of water at 4° C. and under a pressure of one standard atmosphere. Which mass was to be chosen as the standard? Eventually the platinum standard was adopted, so that the kilogram is now defined as the mass of a certain lump of platinum-iridium (a copy of the original Borda kilogramme), and when comparisons with it are being carried out a correction for the buoyancy of the air is to be made if the material of the mass to be compared is not also platinum-iridium [cf. p. 88]. This choice of a unit destroys the simplicity existing in the original definition,

the density of water no longer being 1 gm. cm.$^{-3}$ at 4° C. and under a pressure of one standard atmosphere. The difference is small, but it has to be considered in accurate work dealing with volume determinations. Accordingly, the litre is now defined as follows: 'It is the volume of one kilogram of pure water at the temperature of its maximum density and under a pressure of one standard atmosphere.' On this basis

$$1 \text{ litre} \equiv 1000{\cdot}028 \text{ cm.}^3.$$

The litre and millilitre (ml.) are now frequently chosen as the units of volume, burettes, flasks, etc., being marked in millilitres and fractions thereof. The advantage of this arrangement is that simplicity is regained, for the maximum density of water is 1 gm. ml.$^{-1}$.

The Correction for Stem Exposure.—Let us now investigate the correction to be applied to a mercury thermometer on account of stem exposure. Suppose that t_b is the reading of the thermometer when immersed in a bath whose temperature is t_a, n degree divisions of the thread being exposed. Let t_m be the *mean* temperature of the exposed column as derived from observations on two independent thermometers situated near to it. Now the *volume* corresponding to one degree division may be considered as our unit of volume for this particular purpose. If a is the apparent coefficient of expansion of mercury in glass, viz. 0·00016 deg.$^{-1}$C., then if the exposed column were heated to t_a it would expand $na(t_a - t_m)$, or for practical purposes $na(t_b - t_m)$ since t_a and t_b are nearly equal. The corrected temperature is $t_b + na(t_b - t_m)$. [This calculation is of purely academic interest.]

Example.—$t_b = 250°$ C., $n = 150$, and $t_m = 40°$ C.
$t_a = 250 + (150 \times 0{\cdot}00016 \times 210) = 255{\cdot}0°$ C.

Correction of Barometric Reading for Temperature.—Let H_1 be the height as measured on the scale whose material has a coefficient of linear expansion λ. This is not the true height, since if the scale were graduated at $t_1°$ C. and used at $t_2°$ C., each cm. division which is exact at $t_1°$ will be $[1 + \lambda(t_2 - t_1)]$ cm.

For suppose l_0 is the distance between two consecutive cm. marks when the temperature is 0° C. Then l_1, the distance between these marks at $t_1°$ C., is given by

$$l_1 = l_0(1 + \lambda t_1) = 1 \text{ cm.},$$

since it has been assumed that the scale was constructed at this temperature.

Similarly $$l_2 = l_0(1 + \lambda t_2) \text{ cm.},$$

where l_2 is the distance between the same marks at t_2° C. Hence

$$\frac{l_2}{l_1} = \frac{1 + \lambda t_2}{1 + \lambda t_1}.$$

or $$l_2 = [1 + \lambda(t_2 - t_1)] \text{ cm.}$$

This approximation is justified by the fact that λ is small and t_1 and t_2 are not, in practice, very different from each other.

The barometric reading corrected for the fact that the scale is used at a temperature different from that at which it was made is therefore $H_1[1 + \lambda(t_2 - t_1])) = H_2$ [say]. The pressure is $g\rho_2 H_2$ where ρ_2 is the density at t_2° C.: we require the height H of a column of mercury at 0° C. which would exert this same pressure. If ρ_0 is the density of mercury at 0° C. and a its coefficient of expansion, H is determined by the equation

$$g\rho_0 H = g\rho_2 H_2.$$

Since $$\rho_2 = \frac{\rho_0}{(1 + at_2)},$$

$$H = \frac{H_1[1 + \lambda(t_2 - t_1)]}{[1 + at_2]}.$$

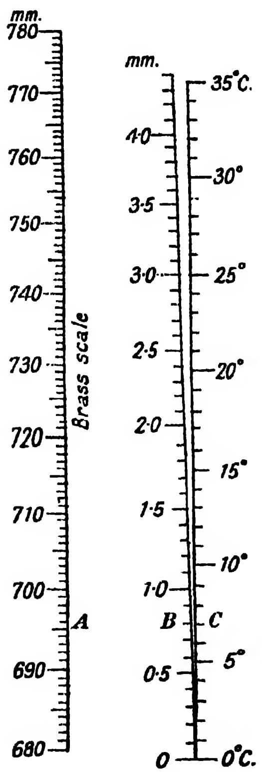


FIG. 9·11.— Mehmke's Method for reducing a Barometric Reading to 0° C. (for scales on brass).

Mehmke's Method for Correcting a Barometric Reading for Temperature.—To determine the corrections to be applied to barometer readings for temperature, assuming the scales to be graduated at 0° C., MEHMKE proceeded as follows. Suppose that the observed reading is 751 mm. of mercury at 17° C. By means of a straight line join these points on the scales A and C shown in Fig. 9·11. The intercept on the scale B—2·1 mm.—then gives the amount to be subtracted from the observed height in order to reduce the reading to 0° C. A device of this sort is known as a *nomograph.*

Gas Regulators and Thermostats.—As the name suggests, a thermostat is a source of constant temperature. If a bath is heated by a gas flame the temperature of the bath is never constant; this is because the supply of gas varies or else draughts exist, and these, being of a variable nature, cause the heated body to lose thermal energy at different rates. The device shown in Fig. 9·12 is used to regulate the supply of gas, so

that when the temperature tends to fall, more gas is supplied, and vice versa. The regulator A is placed inside the bath which is to become the thermostat, and gas entering at D travels along the path indicated by the arrows to the burner. The bulb A contains toluol, this liquid being chosen on account of its high coefficient of expansion but otherwise constant properties. When the temperature rises beyond the desired limit the expansion of the toluol forces the mercury upwards and this seals the tube at B; the screw C is arranged to allow sufficient gas to flow to the burner through E and so prevent complete extinction of the flame. The desired temperature is obtained by altering the position of the narrow tube in B. The rubber at C must be sufficiently long to allow for this manipulation.

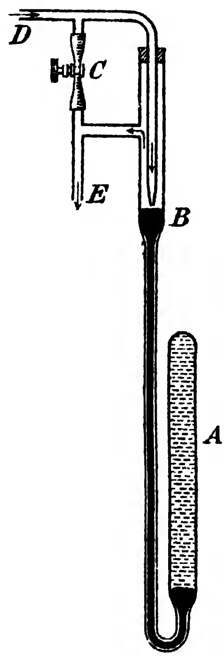
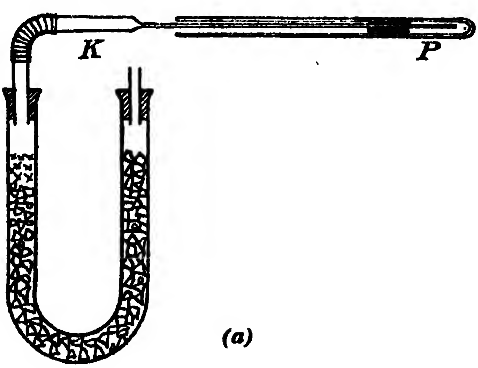


FIG. 9·12.—Gas Regulator for Thermostat.

The Thermal Expansion of a Gas at Constant Pressure.—*Method* (*i*). In an earlier section [cf. p. 85] it was shown that the volume of a given mass of gas at a constant temperature depends upon the pressure to which it is subjected, so that if we wish to investigate how its volume varies with temperature the pressure inside the apparatus must be maintained constant. A convenient form of apparatus and a method of filling it with dry air is indicated in Fig. 9·13. A glass tube of uniform diameter [about 2 mm.] and 40 cm. long having been cleaned, dried, and closed at one end [cf. p. 152], is attached to a scale graduated in cm., etc. To indroduce a pellet of mercury, about 5 cm. long, and situated half-way down the tube, and so enclose a quantity of dry air, an arrangement such as that indicated in Fig. 9·13 (*a*) may be adopted. A quantity of soda lime is placed in a U-tube and one of its limbs is provided with a rubber bung through which passes a glass tube K. This is drawn out to a fine capillary. A gentle stream of air is blown through this apparatus so that the air in K shall be dry. By means of a wide glass tube drawn out to a capillary a pellet of mercury is first introduced into the experimental tube, P. This tube is held in a vertical position and most of the air below the pellet removed by inserting the end of the fine capillary below the pellet. The excess pressure due to the weight of the pellet forces the air below it through the capillary. This is then withdrawn and the tube P placed in a horizontal position with the end of the capillary attached to K projecting beyond the pellet. P is then slightly raised when the pellet

moves slowly down P and dry air is drawn into the tube. This operation is repeated several times to dry the walls of P thoroughly.

When this has been done the tube P is withdrawn and attached in a vertical position to S, Fig. 9·13 (*b*), by means of rubber bands, B and C. This is then placed in a metal container and surrounded by melting ice. The position of P is adjusted so that the upper end of the mercury pellet is visible. The ice is thoroughly stirred and when the position of the pellet becomes constant the temperature of the air in the tube will be 0° C. If the tube is uniform in diameter the length of that portion occupied by the dry air is directly proportional to its volume. To determine this length it is not desirable to raise the tube from the water



(*a*)

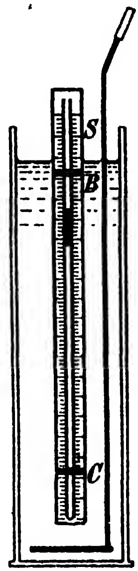


FIG. 9·13.—Simple Air Thermometer. (*b*)

to see the lower end of the pellet since the air inside may be changed in temperature. This difficulty may be avoided by measuring the length of the mercury pellet [the small change in this with temperature being neglected], and observing the position of the upper end of the pellet. In addition, if the scale S extends beyond the open end of the tube, the position of this end should be adjusted to some definite mark on the scale before taking observations, since the tube may move during the course of the experiment.

The length of the tube occupied by dry air having thus been ascertained at 0° C., the ice is removed and the temperature raised to that of steam under the existing atmospheric conditions. This is preferably done by jacketing the tube with a wide brass tube through which steam is passed. The corresponding length of the tube below the pellet is determined.

Now α_p, the coefficient of increase in volume at constant pressure, is defined as the fraction of the volume at the tem-

perature of melting ice by which the volume of a given mass of gas increases for a rise in temperature of one degree, the pressure remaining constant. Hence

$$\alpha_p = \frac{\text{increase in volume at constant pressure}}{\text{volume at the temperature of melting ice} \times \text{change in temperature}}$$
$$= \frac{v_t - v_0}{v_0 \cdot t}.$$

The above experiment enables α_p to be found and it will be noticed that no thermometer has been used. Strictly speaking, α_p as here determined, is an 'apparent' coefficient, but the correction for the expansion of the glass is negligible compared with the experimental errors.

The same apparatus may now be used to determine the difference between a temperature as indicated by a mercury-in-glass thermometer and by a constant-pressure air thermometer. [It should be pointed out that there is no reason at all to suppose that the indications ought to be identical.] Thus, to determine this difference at about 50° C., the tube and mercury thermometer are placed in water and the temperature adjusted by passing in steam [or otherwise]. When the temperature is steady, the reading of the mercury thermometer and the position of the pellet in the tube are noted. Assuming α_p, the temperature of the bath is calculated and the required difference deduced.

GAY LUSSAC, and later REGNAULT, investigated the thermal expansion of gases at constant pressure. They found for the so-called ***permanent*** gases that this coefficient was equal to 0·00367 or $\frac{1}{273}$ deg.$^{-1}$ C. This statement is an expression of a law generally referred to as Charles' Law. [The gases hydrogen, oxygen, nitrogen, helium, are called permanent since at one time it was believed that they could not be liquefied.]

The Pressure Coefficient.—If a gas is heated under the condition that its volume remains constant, the pressure increases.

The coefficient of increase in pressure at constant volume, α_v, is defined as the fraction of the pressure at the temperature of melting ice by which the pressure of a given mass of gas increases for a rise in temperature of one degree, the volume remaining constant. Hence

$$p_t = p_o\,(1 + \alpha_v t).$$

If the gas obeys Boyle's law, it may be shown theoretically that $\alpha_p = \alpha_v$. Let p, V and t be the pressure, volume and temperature of a given mass of gas, whilst suffixes attached to p and V denote the values of these quantities at different temperatures. If the temperature of the gas is increased from 0° to t° while the pressure remains constant,

$$V_t = V_0(1 + \alpha_p t) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

If the temperature of the gas remains at $t°$, but the pressure is increased to p_t until the volume is V_0, then, by Boyle's law,

$$p_0V_t = p_tV_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Eliminating V_t from these equations, we have

$$p_0V_0(1 + \alpha_p t) = p_tV,$$

or

$$p_0(1 + \alpha_p t) = p_t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If, however, the volume had remained constant throughout and the temperature had been increased from the temperature of melting ice to $t°$, then from the definition of α_v

$$p_t = p_0(1 + a_v t). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Hence

$$\alpha_p = \alpha_v.$$

Experimental Determination of α_v, the Pressure Coefficient.—A convenient laboratory method uses the apparatus indicated in Fig. 9·14. A bulb A, containing dry air, is connected by rubber tubing to a mercury reservoir B. The glass tube leading from A passes through a cork in the bottom of a metal vessel D, containing melting ice. C is a fiducial mark to which the level of the mercury is adjusted by raising or lowering B when the temperature of A has been kept constant for several minutes by thoroughly stirring the ice mixture. To determine the difference in height h between the levels of the mercury at C and B, a U-tube is partly filled with water and placed as shown. The required difference is ascertained by means of the scale S. If the barometric height is known the pressure in A may be calculated.

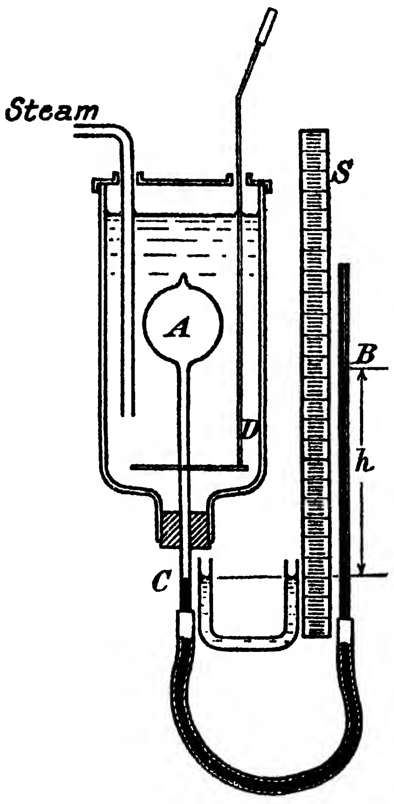


FIG. 9·14.—Apparatus to find a_v.

The ice is then removed from round A and the temperature raised by passing in steam. A finally acquires the steam temperature and after the bulb has

been at this temperature for several minutes and the mercury brought to C, the pressure in the apparatus is determined as before. Then α_v may be calculated, since

$$p_t = p_0(1 + \alpha_v t).$$

Again it should be noticed that α_v has been obtained without reference to a mercury thermometer, and the instrument may be used to compare the corrected reading of a mercury-in-glass thermometer at a steady temperature with the reading of a constant-volume air thermometer at the same temperature: or a curve could be obtained showing the correction to be applied to a given mercury thermometer in order to obtain a temperature on the gas scale.

[*Note:* If the mercury in the tube B is at the temperature of the mercury in the barometer, the height h may be added to the uncorrected height of the barometer, since the final calculation depends on the ratio of pressures. To deduce t, however, the barometer height must be corrected in the usual way.]

The Absolute Zero of Temperature.—Suppose that we try

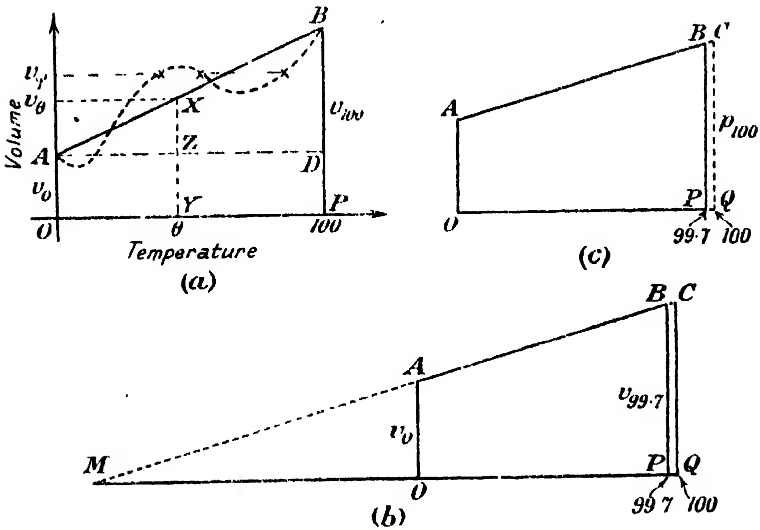


FIG. 9·15.—Centigrade Scale of Temperature and the Absolute Zero of Temperature.

to use a knowledge of the expansion of a gas at constant pressure in order to construct a scale of temperature. Then v_0, the volume of the gas at the temperature of melting ice, and v_{100}, the volume

of the gas under the same pressure but at the temperature of steam produced under standard conditions, have to be determined. If these data have been obtained the diagram shown in Fig. 9·15 (*a*) may be constructed. In this, the straight line OP has been made 100 units long, while at O and at P ordinates to represent v_0 and v_{100} have been drawn—these are OA and PB respectively. In other words, we have plotted the points $(0, v_0)$ and $(100, v_{100})$. Suppose the points A and B are joined by any arbitrary curve such as the one shown dotted. Then this curve could be used to construct a scale of temperature. Thus, if the instrument were used under such conditions that the volume of air increased to a value v_ϕ, the pressure remaining constant, then the temperature would be obtained by drawing a straight line parallel to the temperature axis OP to intersect the curve. For the particular curve drawn there would be three points of intersection—a state of affairs which is utterly absurd since a thermometer cannot be at three different temperatures at one and the same time. This difficulty could be removed by drawing a simpler form of curve so that only one point of intersection would be obtained, but the most logical thing to do is to join A and B by a straight line. Then, in effect, the increment in volume represented by BD has been divided into 100 equal parts, each of which represents a degree on the centigrade scale of a gas-thermometer at constant pressure. Then, as similarly discussed in detail on p. 151, the volume v_θ will represent a temperature θ on this scale. It must be borne in mind that there is no reason why this temperature should be the same as that of a mercury-in-glass thermometer in the same bath as the gas-thermometer—even when all corrections to such a thermometer have been applied.

In actual practice it is seldom possible to obtain v_{100} on account of the atmospheric pressure not being equal to that of a column of mercury 76 cm. long, at sea-level in latitude 45°, when its temperature is that of melting ice. The corrected barometric height is therefore obtained and a value for the steam temperature calculated —for present purposes it is sufficient to assume that the change in the boiling-point of water for one centimetre of mercury change in pressure is the same on all centigrade scales, viz. 0·37°. Thus, suppose the steam temperature is 99·7° C. when the experiment with the gas-thermometer is made. The diagram shown in Fig. 9·15 (*b*) may then be constructed, and if BA is produced to cut PO in M, it will be found that OM represents a temperature of − 273° on the constant-pressure gas centigrade scale of temperature. At this temperature the volume of the gas would be zero if the gas retained its normal properties. It is called the ***absolute zero of temperature*** on the above scale. If T_0 is the temperature of melting ice

on this scale, then $T_0 = 273°$ K.[1] Any temperature $\theta°$ C. is therefore $(T_0 + \theta)°$ K.

The convenience of such an absolute scale is at once apparent for ***the volume of a given mass of gas at constant pressure is directly proportional to its absolute temperature.*** This is known as GAY-LUSSAC's law for a gas at constant pressure. To prove this, we have, $v_\theta = v_0(1 + \alpha_p.\theta)$

But $\alpha_p = \frac{1}{T_0}$, so that

$$v_\theta = v_0\left(1 + \frac{\theta}{T_0}\right) = \frac{v_0}{T_0}(T_0 + \theta).$$

$$\therefore \frac{v_0}{T_0} = \frac{v_\theta}{T_0 + \theta} = \frac{v_\theta}{T_\theta}.$$

Now in a similar way we may use the increase in pressure of a gas heated at constant volume in order to construct another centigrade scale of temperature. Thus if p_0 and $p_{99 \cdot 7}$, say, have been obtained experimentally, the diagram shown in Fig. 9·15 (*c*) may be constructed. If BA is produced it will be found to cut the axis of temperature at $-273°$ on the constant-volume centigrade gas scale. As before, it is easily shown that

$$\frac{p_0}{T_0} = \frac{p_\theta}{T_0 + \theta} = \frac{p_\theta}{T_\theta},$$

i.e. ***the pressure of a given mass of gas at constant volume is directly proportional to its absolute temperature.*** This is GAY-LUSSAC's law for a gas at constant volume.

For an ideal gas, which in the above has been assumed to be the thermometric substance, the two absolute zeros are identical. We shall assume that they are the same for all such gases as air, nitrogen, helium, etc.

The Characteristic Equation for Gases.—If p is the pressure, v the volume, and T the absolute temperature *of* 1 *gm. of gas,* then

$$\frac{pv}{T} = \text{constant} = \mathcal{R}.$$

This equation, which is called the *characteristic equation for a gas*, shows that if T is constant, then pv is constant [Boyle's law]; on the other hand if v is constant $\frac{p}{T}$ is constant, or if p is constant,

$\frac{v}{T}$ is constant [Gay-Lussac's law]. Using c.g.s. units, p is expressed in dyne.cm.$^{-2}$, v has the dimensions cm.^{3} gm.$^{-1}$. Thus pv is

[1] The letter K is used to denote temperatures on an absolute centigrade scale.

expressed in erg.gm.$^{-1}$. Consequently the unit for $\mathcal{R}$ is erg.gm.$^{-1}$ deg.$^{-1}$ K.

In some calculations it is more convenient to consider the volume V of 1 gram-molecule. [A mass of any substance equal to its molecular weight in grams is termed a gram-molecule of that substance.] The characteristic equation then becomes $pV = RT$, where $R = M\mathcal{R}$, M being the molecular weight of the substance. [cf. also p. 109.] R is known as the universal gas constant, whereas $\mathcal{R}$ is the gas constant per gram of gas.

To calculate the value of R, since 1 mole of a gas occupies 22,415 cm.^{3} at S.T.P., we have

$$R = \frac{pV}{T} = \frac{76 \times 13{\cdot}59 \times 980{\cdot}6 \times 22415}{273{\cdot}2}$$

$$\therefore R = 8{\cdot}314 \times 10^7 \text{ erg.mole.}^{-1} \text{ deg.}^{-1}.$$

Hence, for oxygen, molecular weight 32,

$$\mathcal{R}_{O_2} = (8{\cdot}314 \times 10^7) \div 32 = 2{\cdot}598 \times 10^6 \text{ erg.gm.}^{-1}\text{deg.}^{-1}.$$

If V is the volume of a mass m of gas, the characteristic equation becomes

$$\frac{pV}{T} = m\mathcal{R}, \text{ or } \frac{pV}{\mathcal{R}T} = \text{mass of gas (in gm.)}.$$

This equation is utilized in the construction of a standard gas thermometer, but before such an apparatus is described its use in a numerical example will be demonstrated.

***Example.*—Two bulbs of 100 cm.^{3} and 200 cm.^{3} capacity are connected together by means of a capillary of negligible bore. Initially both bulbs are in melting ice; finally the 200 cm.^{3} bulb is in steam at 100° C. If the initial pressure is 76 cm. of mercury, what is the final pressure? [Neglect expansion of the bulbs.]**

Let h be the final pressure (cm. of mercury). Let g and ρ denote the intensity of gravity and the density of mercury respectively.

Consider the *mass* of gas in each bulb.

Initially, $\dfrac{76 \times g \times \rho \times 100}{\mathcal{R} \times 273}$ = mass of gas in the smaller bulb,

and $\dfrac{76 \times g \times \rho \times 200}{\mathcal{R} \times 273}$ = mass of gas in the larger bulb.

$$\therefore \frac{76 \times g \times \rho \times 100}{\mathcal{R} \times 273} + \frac{76 \times g \times \rho \times 200}{\mathcal{R} \times 273} = \text{mass of gas in both bulbs}$$

(and it is this quantity which remains constant).

Similarly, in the final stage

$$\frac{g\rho h \times 100}{\mathcal{R} \times 273} + \frac{g\rho h \times 200}{\mathcal{R} \times 373} = \text{mass of gas in both bulbs}$$

$$= \frac{76 \times g \times \rho \times 100}{\mathcal{R} \times 273} + \frac{76 \times g \times \rho \times 200}{\mathcal{R} \times 273}$$

$\therefore h = 92{\cdot}5$ cm. of mercury.

Boyle's law.—It has already been stated that the volume of a given mass of gas at constant temperature is inversely proportional to the pressure to which it is subjected. In symbols this becomes $pV =$ constant and it is now known that this constant is $m\mathscr{R}T$, where the symbols have their usual significance: thus $pV = m\mathscr{R}T$. It will be noted that this expression involves m, but if ρ is the density of the gas when the pressure is p, so that $\rho = \frac{m}{V}$, we may write $\frac{p}{\rho} = \mathscr{R}T =$ constant, if the temperature is constant. Thus for any ideal gas at constant temperature it has been established that $\frac{p}{\rho} =$ constant, i.e. the density of an ideal gas under such conditions is directly proportional to the pressure: the 'mass' has disappeared from the equation when it is given in this form.

The Constant Volume Gas Thermometer.—A simple form of such a thermometer is shown in Fig. 9·16. A bulb A is connected by means of a capillary tube to a manometer, DE, the space above the mercury being exhausted so that observations on a second barometer are unnecessary. If at any time any gas should find its way into the space above E, it may be forced into the small bulb above the constriction shown by raising F. When F is restored to its normal position a small pellet of mercury remains in the constriction, the gas being entrapped above it. The reservoir F, containing mercury, may be raised or lowered by means of a pulley (not shown in the diagram). To use this apparatus to measure an unknown steady temperature it must first be used to determine the reading on the absolute scale of temperature corresponding to 0° C., the temperature of steam under standard conditions being defined as 100° C. [This is equivalent to finding α_v since the reciprocal of this is the number required.]

To do this the bulb A is first immersed in melting ice and then in steam. In each instance the pressure inside the bulb A when the mercury in the manometer is in contact with the extremity of a piece of black glass, G, which serves as a fiducial mark, is determined.[1]

Simple Theory.—Let us first neglect the volume of the dead space and the expansion of the bulb A. Let $\overline{V}$ be the constant volume of A. Let T_0 be the absolute temperature corresponding to

[1] Attention must be drawn to the fact that the mercury levels in K and F are the same when the mercury level in D is being made to coincide with the tip of G—the tap N being open, of course. If N is closed, as in the diagram, this condition no longer necessarily applies, for F may be lowered without affecting the mercury level in K.

the zero on the centigrade scale. Let t be the steam temperature on the centigrade scale at the time of the experiment. The corresponding absolute temperature is $(T_0 + t)$. Let p_0 and p_t be the pressures in A when its temperature is 0° C. and t° C. respectively.

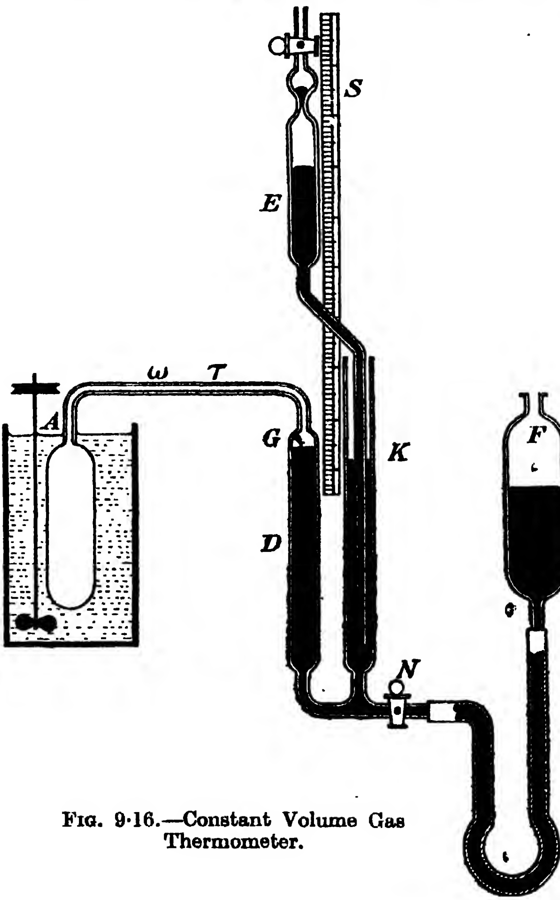


FIG. 9·16.—Constant Volume Gas Thermometer.

Then

$$\frac{p_0\bar{\bar{V}}}{\mathcal{R}T_0} = \text{mass of gas enclosed in A}$$

$$= \frac{p_t\bar{\bar{V}}}{\mathcal{R}(T_0 + t)}.$$

Hence

$$\frac{p_0}{T_0} = \frac{p_t}{T_0 + t}$$

Since p_0, p_t, and t are known, T_0 may be calculated.

Let us now suppose that θ is the temperature on the centigrade

scale which is to be measured. Let p_θ be the pressure in A when its temperature is θ. Then

$$\frac{p_0}{T_0} = \frac{p_\theta}{T_0 + \theta}.$$

so that θ is determinable.

More Complete Theory.—Let p be the pressure and V the volume of A, where suffixes are now used to denote the values of these variables at corresponding temperatures. Let ω be the volume of the dead space, and τ be its mean temperature as measured by calibrated mercury thermometers placed near to it. [Mercury thermometers may be used since terms containing ω and τ only appear as small quantities in the final equations.] Then the symbol τ_θ indicates the mean temperature of ω when A is at $\theta°$ C. Similarly ω_θ is the particular value of ω under these conditions, etc. Then

$$\frac{p_0V_0}{\mathcal{R}T_0} + \frac{p_0\omega_0}{\mathcal{R}(T_0 + \tau_0)} = \text{mass of gas enclosed}$$

$$= \frac{p_tV_t}{\mathcal{R}(T_0 + t)} + \frac{p_t\omega_t}{\mathcal{R}(T_0 + \tau_t)}$$

$$= \frac{p_tV_0[1 + \gamma t]}{\mathcal{R}(T_0 + t)} + \frac{p_t\omega_0[1 + \gamma(\tau_t - \tau_0)]}{\mathcal{R}(T_0 + \tau_t)}$$

where γ is the coefficient of cubical expansion of glass.

To solve this equation for T_0 we first omit all terms containing ω—the 'correction terms'—and use the resulting equation to obtain an approximate value for T_0. This value is then inserted in the correction terms of the more exact equation and the equation thus obtained solved for T_0.

Similarly when A is at $\theta°$ C., we have

$$\frac{p_0V_0}{\mathcal{R}T_0} + \frac{p_0\omega_0}{\mathcal{R}(T_0 + \tau_0)} = \frac{p_\theta V_\theta}{\mathcal{R}(T_0 + \theta)} + \frac{p_\theta\omega_\theta}{\mathcal{R}(T_0 + \tau_\theta)}$$

$$= \frac{p_\theta V_0[1 + \gamma\theta]}{\mathcal{R}(T_0 + \theta)} + \frac{p_\theta\omega_0[1 + \gamma(\tau_\theta - \tau_0)]}{\mathcal{R}(T_0 + \tau_\theta)}$$

The Constant-pressure Gas Thermometer.—The apparatus described on p. 188 may be used as a constant-pressure gas thermometer: in fact, it is a simple form of Gay-Lussac's original air thermometer. Such thermometers are not capable of yielding accurate results, for gas tends to leak passed the mercury pellet. A more modern form is indicated in Fig. 9·17. A bulb S is connected to a mercury manometer by means of a narrow tube AB (1 mm. diameter). The amount of mercury in the manometer, and hence the pressure of the apparatus, may be controlled by the siphon EF. The bulb S and the manometer CE are immersed in baths, the temperature of the former being varied while that of the latter is kept at T_0. Stirrers placed in these baths help to keep the temperatures uniform. Polished metal screens X and Y diminish the exchange of heat between the two sides of the apparatus. To use this thermometer to measure a steady temperature we have to standardize the instrument when its bulb is in ice and then in steam, as in the previous experiment. Let V be the volume of S, ω that of the connecting tubes, and v the volume above the mercury level in the manometer. Let τ be the mean tem-

perature of ω as measured by calibrated mercury thermometers. Then by reasoning similar to that used in the previous section, we have, if Π is the pressure of the gas which is constant throughout the experiment,

$$\frac{\Pi V_0}{T_0} + \frac{\Pi \omega_0}{(T_0 + \tau_0)} + \frac{\Pi v_0}{T_0} = \frac{\Pi V_t}{(T_0 + t)} + \frac{\Pi \omega_t}{(T_0 + \tau_t)} + \frac{\Pi v_t}{T_0},$$

where v_t is the volume of gas above the mercury in the manometer when S is at a temperature t° C.

If γ is the coefficient of cubical expansion of glass, the above equation may be written

$$\frac{V_0}{T_0} + \frac{\omega_0}{(T_0 + \tau_0)} + \frac{v_0}{T_0} = \frac{V_0(1 + \gamma t)}{(T_0 + t)} + \frac{\omega_0[1 + \gamma(\tau_t - \tau_0)]}{(T_0 + \tau_t)} + \frac{v}{T_0}.$$

Hence T_0 may be calculated, since v_t is known from the position of the mercury in the manometer.

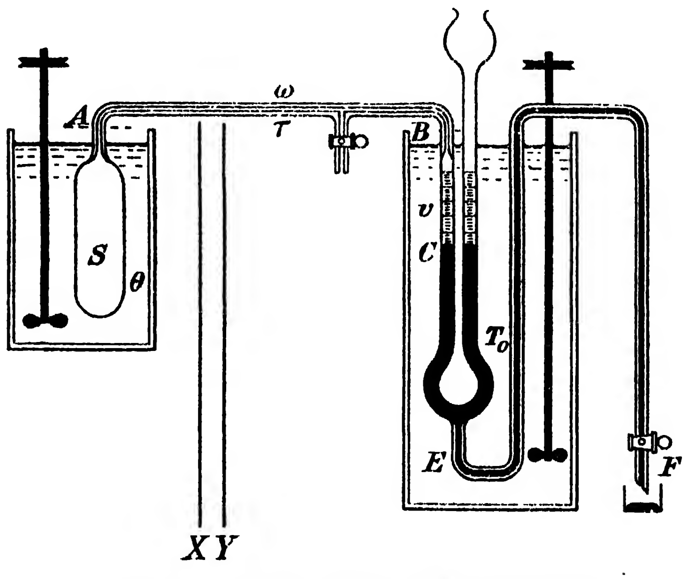


FIG. 9·17.—Constant Pressure Gas Thermometer.

Similarly

$$\frac{V_0}{T_0} + \frac{\omega_0}{(T_0 + \tau_0)} + \frac{v_0}{T_0} = \frac{V_0(1 + \gamma \theta)}{T_0 + \theta} + \frac{\omega_0[1 + \gamma(\tau_\theta - \tau_0)]}{(T_0 + \tau_\theta)} + \frac{v_\theta}{T_0},$$

so that θ may be determined since T_0 is known.

This apparatus may be used to determine the expansion of an irregular solid or the change in volume occurring when a metal melts. If W is the volume of the metal in S the volume of gas is V − W, so that the general equation becomes

$$\frac{V_\theta - W_\theta}{(T_0 + \theta)} + \frac{\omega_\theta}{T_0 + \tau_\theta} + \frac{v_\theta}{T_0} = \text{constant.}$$

If the volume of the metal in the apparatus is known at one temperature, the above equation enables us to deduce that at a second temperature when the experiment has been carried out at these two temperatures.

Callendar's Compensated Gas Thermometer (Constant Pressure).—The gas thermometer described in the previous paragraph is not a precision instrument, chiefly owing to the errors arising in the determination of v, since the mercury surface has to be viewed through a water bath. To avoid this and also eliminate the correction for dead space CALLENDAR devised the compensated gas thermometer shown diagrammatically in Fig. 9·18. The 'thermometer side' of the instrument consists of a glass [or silica] bulb V attached by capillary tubing 1 mm. in diameter to another bulb M containing mercury.

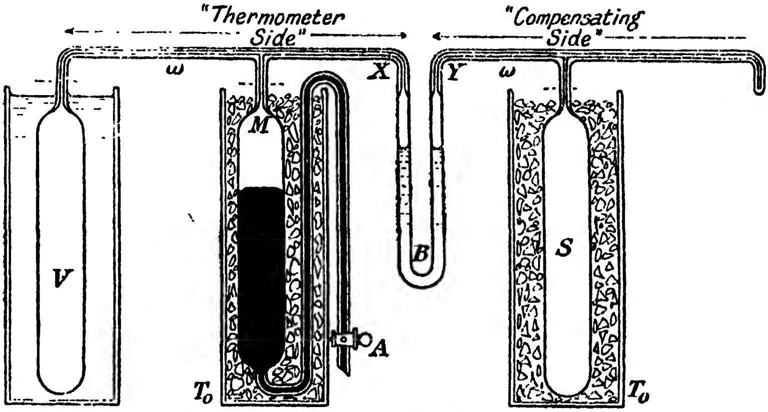


FIG. 9·18.—Callendar's Compensated Gas Thermometer.

This is kept in melting ice, and mercury may be withdrawn from it through the siphon A. The 'compensating side' of the instrument consists of a bulb S maintained at 0° C. throughout the experiment. These two parts are connected together by an oil gauge B.

Actually the apparatus was bent at the points X and Y so that the capillary tubes lay side by side and were therefore at the same mean temperature. Initially all bulbs are at 0° C. and the pressure, Π, and mass, m, of gas on each side of the gauge made equal. Let V, S and M be the volumes of the gas in V, S, and above the mercury in M respectively. Hence, when all three bulbs are in melting ice, we have

$$\frac{\Pi V_0}{\mathscr{R}T_0} + \frac{\Pi\omega_0}{\mathscr{R}(T_0 + \tau_0)} + \frac{\Pi M_0}{\mathscr{R}T_0} = \frac{\Pi\omega_0}{\mathscr{R}(T_0 + \tau_0)} + \frac{\Pi S_0}{\mathscr{R}T_0} = m,$$

$$\text{i.e. } \frac{V_0}{T_0} + \frac{M_0}{T_0} = \frac{S_0}{T_0}.$$

Similarly, when V is at t° C.,

$$\frac{\Pi V_t}{\mathscr{R}(T_0 + t)} + \frac{\Pi\omega_t}{\mathscr{R}(T_0 + \tau_t)} + \frac{\Pi M_t}{\mathscr{R}T_0} = \frac{\Pi\omega_t}{\mathscr{R}(T_0 + \tau_t)} + \frac{\Pi S_0}{\mathscr{R}T_0} = m,$$

$$\text{i.e. } \frac{V_t}{(T_0 + t)} + \frac{M_t}{T_0} = \frac{S_0}{T_0}.$$

Hence

$$\frac{V_0}{T_0} + \frac{M_0}{T_0} = \frac{V_t}{(T_0 + t)} + \frac{M_t}{T_0},$$

or

$$\frac{V_0(1 + \gamma t)}{T_0 + t} = \frac{V_0 + M_0 - M_t}{T_0},$$

so that T_0 is determinable.

Similarly, when V is at $\theta°$ C.,

$$\frac{V_0(1 + \gamma\theta)}{T_0 + \theta} = \frac{V_0 + M_0 - M_\theta}{T_0},$$

so that θ is known.

If the bulb is made of silica, γ may be neglected, and we have

$$\frac{T_0 + \theta}{T_0} = \frac{V_0}{V_0 + M_0 - M_\theta}$$

i.e.

$$\theta = T_0\left[\frac{M_\theta - M_0}{V_0 + M_0 - M_\theta}\right]$$

EXAMPLES IX

1.—Mercury has a density of 13·59 gm. cm.$^{-3}$ at 0° C. What is its density when placed in steam (barometer 75·1 cm. of mercury).

2.—A glass weight thermometer has a mass 6·34 gm. when empty, and 151·73 gm. when filled with mercury at 99° C. If 2·08 gm. have been expelled in changing the temperature from 0° C. to 99° C., determine the coefficient of relative expansion for mercury in glass.

3.—A mercury thermometer has a stem 18·6 cm. long, the internal diameter of which is 0·068 cm. The thermometer is to be used from − 5° C. to 110° C. Calculate the maximum volume of the bulb if the expansion of mercury in glass is 0·00015. deg.$^{-1}$ C.

4.—Convert the following values to S.T.P.

(*a*) 250 cm.^{3}, 17° C., 78 cm. pressure.
(*b*) 1092 cm.^{3}, 101° C., 40 cm. pressure.

5.—A litre of air at S.T.P. (i.e. standard temperature and pressure, 0° C. and 76 cm. of mercury at 0° C. etc. respectively) has a mass of 1·29 gm. At what temperature will the mass of a litre of air be unity under a pressure of 768 mm. of mercury ?

6.—A sample of dry gas is contained in two vessels connected together by a tube of negligible volume. Both vessels are initially at 20° C. One is raised to 100° C. What is the final pressure if the initial pressure is 72 cm. of mercury, and each bulb has a constant volume of 48·6 c.c. ?

7.—Describe a modern form of gas thermometer and explain how you would use it to measure an unknown but steady temperature.

8.—Describe an accurate method of determining the coefficient of absolute expansion of mercury. If a column of liquid 50 cm. long at 4° C. balances a column of the same liquid 50·5 cm. long at 98° C. calculate the absolute coefficient of cubical expansion of the liquid.

9.—The density of mercury at 10° C. is 13·57 gm. cm.$^{-3}$. At 100° C. it is 13·35 gm. cm.$^{-3}$. What is the mean coefficient of expansion of mercury between these two temperatures ?

10.—A weight thermometer of fused silica holds 5,000 gm. of a liquid at 25° C. When heated to 100° C., 0·400 gm. of liquid overflows. Assuming that the liquid expands uniformly how much more liquid will overflow when the temperature is raised to 175° C.? The expansion of the silica may be neglected. Justify any formula you quote.

11.—Distinguish between the absolute and apparent coefficients of expansion of mercury, and explain how the former coefficient has been directly determined.

12.—Explain how (*a*) the change of volume of a gas heated under constant pressure, (*b*) the change of pressure of a gas heated at constant volume, may be used to define a scale of temperature. Show that if the gas is an ideal gas the two scales will agree. Describe a form of apparatus suitable for measuring temperatures on the first scale.

13.—In 1802 Dalton observed that 1,000 volumes of air at 55° F. become 1,321 volumes at 212° F., the pressure being constant. Compare the value of the coefficient of expansion of air at constant pressure given by these observations with the ordinary text-book value.—(N.H.S.C. 29.)

14.—Show that in the case of a water-in-glass dilatometer the temperature of apparent maximum density will be above the true temperature, and find how much mercury must be placed in the bulb of such a dilatometer in order that the temperature of apparent maximum density may be identical with the true temperature, assuming that the coefficients of volume expansion of mercury and glass are 0·000180 deg.$^{-1}$ C. and 0·000023 deg.$^{-1}$ C. respectively.

15.—Describe experiments to show that water has a temperature of maximum density at about 4°.

Discuss the bearing of this fact on the freezing of water in lakes.

16.—A cylinder of iron 30 cm. long floats vertically in mercury at 0° C. Calculate the increase in the depth to which the cylinder sinks when the temperature is raised to 100° C. Density of mercury at 0° C. = 13·6 gm. cm.$^{-3}$; density of iron at 0° C. = 7·6 gm. cm.$^{-3}$; absolute coefficient of expansion of mercury = $1{\cdot}82 \times 10^{-4}$ deg.$^{-1}$ C.; coefficient of cubical expansion of iron = $3{\cdot}51 \times 10^{-5}$ deg.$^{-1}$ C.

17.—An iron bottle holds 1,380 gm. of mercury when full at 0° C., 20 gm. of mercury are expelled when the bottle is heated to 100° C. and a further 42 gm. when it is heated in an oven. Assuming that the expansion of the bottle may be neglected, calculate a value for the temperature of the oven.

18.—Distinguish between the *true* and *apparent* coefficients of expansion of a liquid, and explain how the true coefficient of expansion of mercury has been *directly* determined.

19.—Describe how you would determine the coefficient of expansion of a liquid by floating a Nicholson's hydrometer of known expansibility in it. If a liquid has a density of 0·831 gm. cm.$^{-3}$ at 15° C. and 0·793 gm. cm.$^{-3}$ at 82° C., calculate its mean coefficient of expansion between these temperatures.

20.—Give the theory of a weight thermometer and describe how you would use it to determine the coefficient of expansion of mercury.

21.—A barometer reads 754 mm. at 17° C. Find the reading at 0° C., if the apparent coefficient of expansion of mercury in glass is 0·00016 deg.$^{-1}$ C., and the coefficient of linear expansion of glass is 9×10^{-6} deg.$^{-1}$ C. Assume that the instrument is furnished with a glass scale correct at 0° C.

22.—The density of water at 4° C. is unity, and at 60° C. 0·983 gm. cm.$^{-3}$. Calculate the mean coefficient of expansion of water between these two temperatures.

23.—Describe Regnault's method of determining the coefficient of expansion of mercury.

24.—A flask containing dry air is corked up at 20° C., the pressure being 1 atmosphere. Calculate the temperature at which the cork will be blown out if this occurs when the pressure inside the flask is 1·7 atmospheres.

25.—In an experiment to determine the expansion of a liquid by the method of Dulong and Petit, the heights of the columns were 59·72 cm. and 61·08 cm., the temperatures being 0·0° C. and 99·7° C. respectively. Calculate the coefficient of expansion of the liquid.

CHAPTER X

CALORIMETRY

Quantity of Heat.—The fact that two or more equal masses of different materials are at the same temperature does not mean that they are thermally alike, i.e. if they are placed into equal quantities of water at the same initial temperature, the rise, or fall, in temperature of the water will be different. This is because the bodies contain different quantities of thermal energy. The unit quantity of heat is called the ***calorie*** and it is the amount of heat required to raise the temperature of one gram of water 1° C. This particular unit of heat is sometimes called the ***small*** or ***gram-calorie*** to distinguish it from the ***large*** or ***kilogram-calorie,*** which is defined as the amount of heat necessary to raise the temperature of one kilogram of water 1° C. Accurate experiments have shown that the amount of heat (energy) required to raise one gram of water 1° C. depends upon the particular degree interval chosen. In practice it is customary to use the ***mean calorie*** which is defined as the hundredth part of the heat required to raise one gram of water from 0° C. to 100° C. For accurate scientific work, where this variation in the calorie has to be considered, it is preferable to define the gram-calorie as one-fifth of the heat necessary to raise the temperature of one gram of water from 15° C. to 20° C. The reasons for this are twofold—(*a*) this is the range in temperature in which experiments are usually made, (*b*) a rise in temperature of about 5° C. is about the smallest range in temperature which can be measured accurately, and it is fallacious to base a practical science on a unit, which cannot be measured with the precision required by modern physics.

Engineers use another unit of heat known as the ***British thermal unit*** [B.T.U.]; it is equal to the heat required to raise the temperature of 1 lb. of water 1° F. The heat necessary to raise the temperature of 1 lb. of water 1° C. is sometimes used in English-speaking countries. It is termed the ***centigrade heat unit***. Gas engineers find these units too small for their requirements so that they have adopted as their unit of heat the ***therm***. It is equal to 100,000 B.T.U.

Thermal Capacity: Specific Heat.—The amount of heat, Q, required to increase the temperature of a body by θ, is proportional to θ if we assume that the physical properties of the body remain constant. Thus

$$Q \propto \theta.$$

To get rid of the sign of proportionality we introduce a constant C, and write

$$Q = C\theta.$$

C is known as the ***thermal capacity*** of the body. If Q is measured in calories, we have

$$[\text{cal.}] = [\text{C}].[\text{deg. C.}]$$

$$\therefore [\text{C}] = [\text{cal. deg.}^{-1}\text{ C.}]$$

Now the thermal capacity of a homogeneous body is proportional to its mass, m, for if its mass is increased n-fold, the heat required to change its temperature by an amount θ will be increased n-fold. Thus

$$C \propto m,$$

or

$$C = sm,$$

where s is a constant known as the ***specific heat*** of the material of the body. The specific heat of a substance is its chemical capacity per unit mass, so that

$$[s] = \text{cal. gm.}^{-1}\text{ deg.}^{-1}\text{ C.}$$

[In another system of units we have, for example,

$$[s] = \text{B.T.U. lb.}^{-1}\text{ deg.}^{-1}\text{ F.}]$$

If M is the mass of water having a thermal capacity equal to that of a given body, mass m and specific heat s, then

$$M \times 1 = ms.$$

M is known as the ***water equivalent*** of the body.

Determination of Specific Heats by the Method of Mixtures.—The following example will perhaps illustrate this method before we discuss it in detail:—

Example.—A block of tin, mass 502 gm., was heated in boiling water at 99·6° C. and then dropped into 313 gm. of water; the temperature rose from 15·4° C. to 19·1° C. Find the specific heat, s, of the tin.

We assume that all the heat given out by the tin in cooling from 99·6° C. to 19·1° C. is acquired by the water.

Now heat lost by tin = mass of tin × its specific heat × its fall in temperature

$$= 502 \times s \times (99{\cdot}6 - 19{\cdot}1) \text{ cal.}$$

Similarly, heat gained by water $= 313 \times 1 \times (19{\cdot}1 - 15{\cdot}4)$ cal.

Equating these two quantities

$$s = 0{\cdot}029 \text{ cal. gm.}^{-1}\text{ deg.}^{-1}\text{ C.}$$

Regnault's Apparatus for Determining Specific Heats.— In an accurate determination of specific heats several errors in the above experiment have to be eliminated. (α) While the metal is being transferred from the heater to the water some heat is lost; (β) we have neglected the heat given to the calorimeter, i.e. the vessel containing the water; (γ) directly the temperature of the calorimeter differs from that of its surroundings there is an exchange of heat between them. To reduce the magnitude of the error due to (α) Regnault devised the apparatus shown in Fig. 10·1. The substance is suspended by means of a piece of cotton inside the heater through which steam is passed. A thermometer is inserted so that its bulb is in contact with the solid whose specific heat is being determined—not to measure the *steam* temperature—but merely to indicate when the temperature of the solid has become

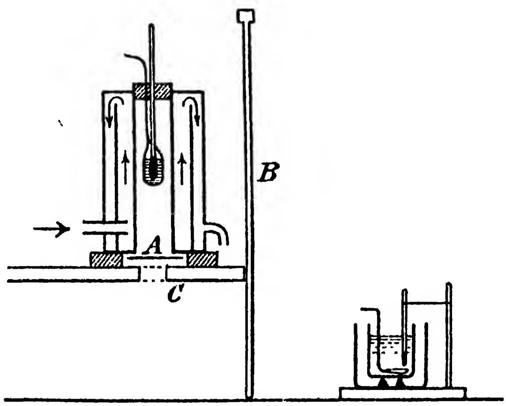


FIG. 10·1.—Regnault's Specific Heat Apparatus for Solids and Liquids.

constant. When steady conditions have been obtained, the screen B is raised, the calorimeter pushed underneath the heater and the solid introduced into the calorimeter by withdrawing the slide A. The calorimeter is quickly withdrawn, the screen lowered, and the rise in temperature of the calorimeter and its contents observed. The specific heat of the solid can be calculated from the following equation :—

Heat lost by solid = heat gained by water + heat gained by calorimeter,

i.e.

(mass of solid × its sp. ht. × its fall in temp.) = (mass of water × 1 × its rise in temperature) + (mass of calorimeter × its sp. ht. × its rise in temperature).

In this equation allowance has been made for the heat imparted to the calorimeter.

Two methods of obtaining a correction for the heat exchange between a calorimeter and its surroundings are discussed below.

Although water is generally used as the calorimetric liquid, it has been suggested that aniline would be better for two reasons, (α) its specific heat is 0·62 cal. gm.$^{-1}$ deg.$^{-1}$ C., so that the rise in temperature is greater than with water when equal amounts of heat are received by the same mass of the two liquids, (β) its vapour pressure is less, so that losses due to evaporation are reduced.

The Specific Heats of Liquids.—These may be determined by the above method if a solid of known specific heat is used in the heater and a known mass of liquid is placed in the calorimeter. The equation to be used is

(mass of solid $\times$ its sp. ht. $\times$ its fall in temperature)

= [(mass of calorimeter $\times$ its sp. ht.) + (mass of liquid $\times$ its sp. ht.)] $\times$ (rise in temperature).

Methods of Calculating the Correction for Heat Exchange between a Calorimeter and its Surroundings in Calorimetric Experiments.—Since, in general, in a calorimetric experiment the temperature of the calorimeter is not the same as that of its surroundings, there must be a heat exchange between them. The rise in temperature if the heat exchange were zero may be obtained as follows :—

(i) ***Rumford's method.***—Rumford first made a correction for this in the following way. By means of a preliminary experiment he ascertained approximately what the rise in temperature was in a given experiment. Let this rise be $\theta°$. He then repeated the experiment with the initial temperature of the calorimeter and its contents $\frac{1}{2}\theta°$ below the temperature, $t°$, of the surroundings. The maximum temperature reached in the repeated experiment will be $(t + \frac{1}{2}\theta)°$ approximately—actually $(t + \phi)°$—and Rumford expressed the view that the heat gained by the calorimeter during the time that the temperature was below t will be compensated by the heat loss when its temperature is above t. This would only be true strictly if the rate of supply of heat to the calorimeter, etc., was constant : in general, this is not so, for the rate of supply diminishes rapidly when equilibrium of temperature between the 'hot body' and the 'calorimeter' is nearly reached—for example, it may happen that the temperature changes from $(t - \frac{1}{2}\theta)°$ to $t°$ in a time which is only one-quarter that in which the temperature changes from $t°$ to $(t + \phi)°$.

(ii) ***Ferry's method.***—This is a simple modification of an

earlier method due to Rowland. In this, readings of the temperature of the calorimeter and its contents are recorded at known times, both before and after the introduction of the hot body, and also during the interval in which the temperature of the hot substance is becoming equal to that of the calorimeter. Suppose that ABCD, Fig. 10·2, is the curve representing such a series of readings. Let the straight line $y = t$ (the room temperature) intersect the above curve in P. From B to P the calorimeter and its contents receive heat from the body which has been introduced into it and also from its surroundings: from P to C they continue to receive heat from the body, but impart heat to the surroundings.

Through P let a straight line parallel to the axis Oy be drawn, and let DC and AB be produced to cut this line in Q and R respectively. Then, in the absence of heat exchange between the calorimeter and its surroundings, the rise in temperature would be RQ. The justification for this is as follows:—If the hot body had not been introduced into the calorimeter, the temperature of the latter would continue to change along BR. Thus, while the temperature actually changes from B to P, the change from M to R was due to heat received from the surroundings, while the change from R to P was due to heat received from the hot body in the time BM.

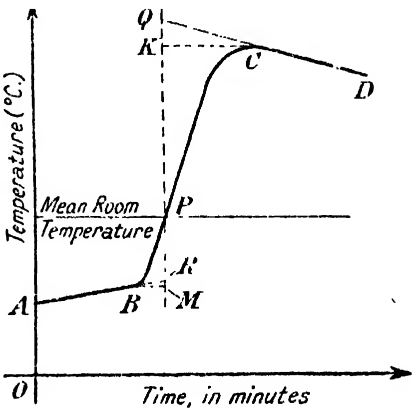


FIG. 10·2.—Forry's Method for Correcting for Heat Exchange between a Calorimeter and its Surroundings.

Now suppose that when the temperature of the calorimeter and the calorimetric liquid is equal to that of the room, i.e. as represented by P, an amount of heat sufficient to raise the temperature by PQ instantaneously is added. Then in the time KC during which the calorimeter, etc., actually continue to receive heat, the loss of heat to the surroundings must be such that the change in temperature is QK, i.e. the point C is reached either along the path PQC or by the actual path PC. Hence QR is the required corrected rise in temperature.

Steady Flow Calorimetry.—This method of determining the specific heat of a liquid [or gas], originally developed by CALLENDAR, is suitable not only for finding the specific heat of a liquid at room temperatures but also at other temperatures—or rather

the actual quantity which is measured is the mean value of the specific heat over a small range of temperature [cf. the definition of the calorie, p. 203].

An apparatus suitable for laboratory work is shown in Fig. 10·3. A narrow glass tube, BC (2 mm. in diameter), is attached to two wider glass tubes, D and E, the whole being supported by means of ebonite discs fitted in a glass tube, FG.

A manganin wire passes down the tube BC and is insulated thermally from the glass by means of a thin rubber cord wrapped spirally round the wire. In this way uniformity of temperature over any cross-section of the tube is secured in the experiment.

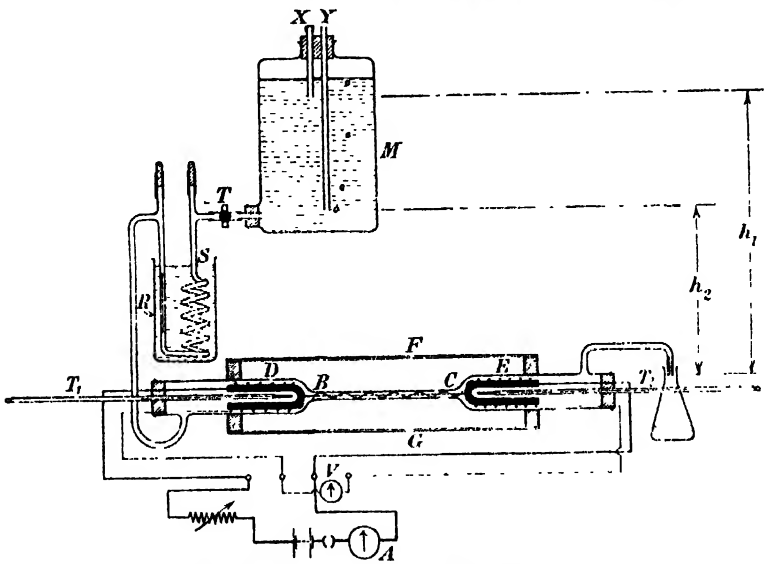


FIG. 10·3.—Steady Flow Calorimeter.

The ends of the wire are soldered to copper cups. The heating current is supplied from a large battery so that the current shall be steady, and it is measured by an ammeter, A. The current leads are shown thick. To determine the rate at which energy is dissipated in the wire and in this instance transferred to the liquid flowing through the capillary tube, it is necessary to measure the potential difference across the wire as well as the current through it. This is done by means of the potential leads (shown thin) and the voltmeter, V. The potential leads are soldered to the copper cups at points near to the ends of the manganin wire. They are wrapped spirally round the cups after being threaded through rubber tubing of such diameter that any liquid flowing through the calorimeter circulates round the cups. In this way uniformity of temperature

is secured across the bulbs of the mercury thermometers, T_1 and T_2, used to measure the rise in temperature of the liquid passing over the manganin wire when energy is dissipated therein, i.e. the thermometers do indicate the temperatures of the incoming and escaping liquid. This arrangement is a very essential part of the apparatus.

It is also necessary to maintain a steady flow of liquid. The Mariotte bottle, M, contains a supply of air-free liquid. It escapes from the bottle when the tap, T, is opened, and passes through a spiral tube, S, immersed in a large beaker of water, so that its temperature becomes constant before it passes through the capillary tube, R, also immersed in the water. The diameter and length of this tube must be so chosen that the liquid flows through the calorimeter at a convenient rate. At a given temperature the rate at which a liquid flows through a particular capillary tube is determined by the pressure difference between the ends of the tube. To secure a constant head of liquid the bottle, M, is provided with two tubes, X and Y, passing through a rubber bung which fits the neck of the bottle tightly. Let us suppose that the shorter tube is used first. When T is opened, the whole of the calorimeter having been filled with the liquid whose specific heat is required, liquid begins to flow through the calorimeter. The flow is not steady until bubbles of air appear at the end of the shorter capillary tube X. The pressure is then determined by the head of liquid h_1. The liquid above the lower end of the tube X is the real supply. In the course of the experiment it will be found necessary to decrease the flow of liquid to about one-half that used at first; this is done by closing the open end of X: bubbles of air soon appear at the lower end of Y, when the pressure is determined by the head of liquid h_2. The exit tube from the calorimeter ends in a short piece of capillary tubing pointing vertically downwards so that the liquid breaks away from the tube in drops.

If I is the current in amperes through the wire and V the potential difference in volts across it, then the rate at which energy is being dissipated is VI watt. [1 watt. $=$ 1 joule. sec.$^{-1}$ $=$ 10^7 erg. sec.$^{-1}$] If m is the mass of liquid flowing per second, s its specific heat, θ the rise in temperature, it is known that

$$VI = Jms\theta,$$

where J is a constant numerically equal to 4·184 in the above system of units. The above equation enables the specific heat of the liquid to be determined. [J $=$ mechanical equivalent of heat in joule. cal.$^{-1}$, cf. p. 272.]

It will be noticed that there is no term in the above equation which takes into consideration the thermal capacity of the calorimeter. In this particular type of calorimetry the thermal capacity of the

calorimeter does not have to be considered since its temperature assumes a steady value at all points before any measurements are made, and therefore no more heat is given to it. It must be pointed out, however, that the method is only suitable for research work when due precautions to obtain steady conditions are taken. If conditions are not kept steady the correction for the water equivalent of the calorimeter is indeterminate and accurate results are not possible.

In the above equation we have assumed that the heat lost is zero. Actually this is not so, although in the apparatus used by Callendar FG was exhausted to reduce heat losses due to conduction and convection, and silvered to diminish the loss of heat by radiation. The correction may be made as follows :—Let h cal. be the heat lost per second per degree rise in temperature above that of the outer jacket. [In Callendar's apparatus FG was surrounded by a second jacket containing water at the temperature of the incoming liquid, in order that the conditions under which heat is lost should be constant.] Then the more exact equation is

$$V_1I_1 = J[m_1s\theta + h\theta],$$

where the suffix denotes one particular experiment. The flow of liquid is then altered to m_2 and the current adjusted so that the rise in temperature of the liquid is again θ. Then

$$V_2I_2 = J[m_2s\theta + h\theta].$$

By subtracting these equations we obtain

$$V_1I_1 - V_2I_2 = Js\theta(m_1 - m_2),$$

which is an equation independent of h, the heat loss as defined above.

In the above we have used a voltmeter and an ammeter to measure the potential difference and current respectively. In the original research these were determined with the aid of a carefully calibrated potentiometer. Moreover, the temperatures were recorded by platinum thermometers.

The Determination of Specific Heats by the Method of Cooling.—***Theory of the Method:*** This method is based on the assumption that when a body cools, while suspended in an enclosure, the quantity of heat δQ, emitted in a time δt, depends only on θ, the excess of the temperature of the body above that of the enclosure, and on the nature and area of the surface of the body. We may therefore write

$$\delta Q = a \, . f(\theta) \, . \delta t$$

where $f(\theta)$ is an unknown function of θ, but one which depends on θ only; a is a constant for any given body.

Let m_1 be the mass of the body, s_1 the specific heat of its material, and $\delta\theta_1$ the fall in temperature in time δt.

Then

$$m_1s_1 \,.\, \delta\theta_1 = \delta Q = a_1 f(\theta)\delta t.$$

$$\therefore\ m_1s_1\frac{\delta\theta_1}{\delta t} = a_1 f(\theta).$$

$$\therefore\ m_1s_1\alpha_1 = a_1 f(\theta),$$

[where $\alpha_1 = \left(\frac{\delta\theta_1}{\delta t}\right)_{\delta t \to 0}$, i.e. α_1 is the drop in temperature per unit time.]

Similarly $$m_2s_2\alpha_2 = a_2 f(\theta).$$

$$\therefore\ \frac{m_1s_1}{m_2s_2} = \frac{\alpha_2}{\alpha_1}, \ldots \text{ if the surfaces are identical in all respects.}$$

From the above equation we see that if the specific heat of the material of one body is known that of another may be determined, when the rates of cooling (i.e. the rates of drop in temperature) for the same excess temperature are known. Moreover, theoretically, the method enables us to determine the specific heat of a substance at a given temperature, instead of a mean value for the specific heat over a range of temperature. In practice, as the method is ordinarily carried out, the inherent errors are often greater than any variation in the specific heat of the material.

Application to Liquids.—In practice the above method is only used for liquids since it is impossible to ensure that the temperature of a solid is uniform throughout. Unfortunately, however, the liquid must be placed in a container and the thermal capacity of this enters into the equation. Thus, if m is the mass of the container and s the specific heat of its material, M_1 the mass of liquid, and S_1 its specific heat, we have

$$(ms + M_1S_1)\delta\theta_1 = \delta Q = a_1 f(\theta) \,.\, \delta t$$

and, similarly, for a second liquid

$$(ms + M_2S_2)\delta\theta_2 = a_2 f(\theta) \,.\, \delta t.$$

Now by using the ***same volume*** of liquid in each instance, and liquids successively in the same container, we may assume (at least as a first approximation) that $a_1 = a_2$. We then have

$$\frac{ms + M_1S_1}{ms + M_2S_2} = \frac{\alpha_2}{\alpha_1},$$

so that if s and S_1 are known (water is generally used as one of the liquids), S_2, the specific heat of the liquid under investigation, may be calculated.

Practical Details.—A blackened calorimeter is about two-thirds filled with water at 70° C., and the temperature noted at half-minute intervals. Since it is necessary that the temperature shown by the thermometer should also be that of the walls of the calorimeter the water must be stirred, but not vigorously. Also,

the calorimeter must be provided with a lid to prevent evaporation of the liquid and consequent loss of heat, which would be a considerable fraction of the heat lost during an experiment by the calorimeter and its contents. For convenience the lid may support the thermometer. To obtain good results, the calorimeter should be suspended by means of three fine strings in a double-walled vessel, the space between the walls being filled with water—Fig. 10·4 (*a*). The temperature of this is recorded—it is the temperature of the enclosure. From the results thus obtained a cooling curve is constructed—Fig. 10·4 (*b*). The problem then immediately before us is to determine the slope of this curve at a given point (temperature). To do this the tangent at any point, P, say, is drawn, its position being estimated by eye. The slope of this tangent determines the rate of cooling for the particular instant represented by P. Since it is impossible to estimate the position of the tangent accurately in this way, let us see how the value for the slope thus obtained may be improved. The process is repeated for several other temperatures and a graph showing the relation between excess temperature and the rate of cooling drawn—cf. Fig. 10·4 (*c*). The particular value of the rate of cooling for a given temperature excess may be deduced from the graph. The value so obtained will be better than that obtained by drawing the tangent at the corresponding point on the cooling curve shown in Fig. 10·4 (*b*), since the position of a tangent can only be estimated, whereas the value for the slope now obtained is derived from the positions of several tangents and we assume that the errors in drawing these are as often positive as they are negative.

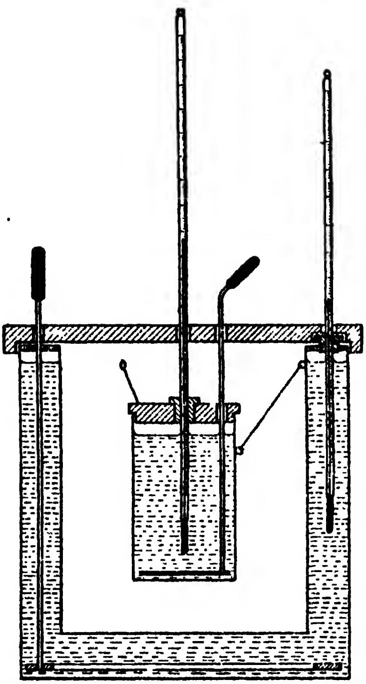


Fig. 10·4 (*a*).—Specific Heat of a Liquid by the Method of Cooling.

An equal ***volume*** of liquid whose specific heat is to be investigated, having been warmed, is then introduced into the calorimeter, and a cooling curve obtained as before. The rate of cooling for the same excess temperature is then determined. From the

information thus obtained, by means of the above equation, the specific heat of the liquid is calculated.

[It must be very carefully noted that the validity of Newton's

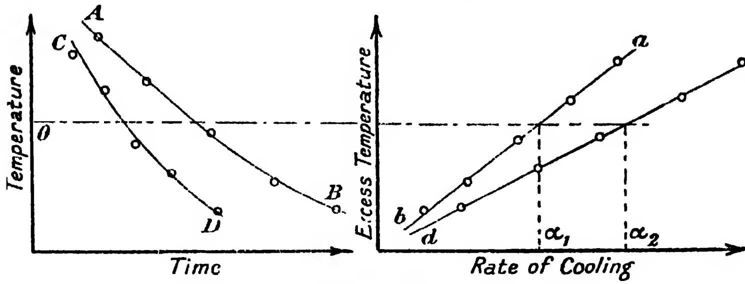


FIG. 10·4.

law of cooling, so often, yet wrongly, connected with the above method of determining specific heats, does not enter at all into the argument.]

On Drawing Tangents to a Curve.—Suppose it is required to draw a tangent at a point P on a curve APB, Fig. 10·4 (*d*). A small plane mirror MN, with its plane normal to that of the diagram, is placed across the curve at P as indicated. PB′ is the image of PB formed by reflexion in the mirror, and the curve appears to be broken at P. If the mirror is slowly rotated about a vertical axis through P, a position of the mirror is found for which the image of PB appears to be continuous with PB itself. The mirror is then normal to the curve at P, so that the tangent is readily constructed.

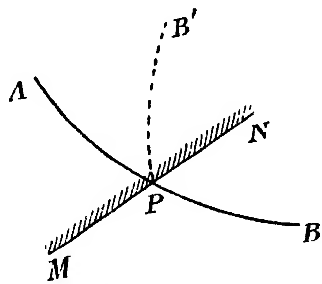


FIG. 10·4 (*d*).

Regnault's Experiments on the Method of Cooling.—REGNAULT made a series of experiments with the object of ascertaining how far this method could be relied upon in the estimation of specific heats. He worked with substances whose specific heats had been determined by the method of mixtures. Regnault showed that the method was not applicable to solids, for not only did his results differ from those obtained by the method of mixtures, but they were not consistent among themselves. A similar conclusion was arrived at with regard to powders. For liquids, however, he concluded that the method was convenient and accurate.

Callendar's Remarks on the Method of Cooling for Liquids.—The advantage of this method is that there is no mixing of the substances under investigation, and consequently no heat due to chemical

action evolved or absorbed. The defect, however, lies in the fact that the whole measurement depends on the assumption that the rate of loss of heat is the same in the two experiments under conditions apparently similar, and that any variation in the conditions or uncertainty with respect to the rate of loss of heat, produces its full effect in the final result, whereas in the method of mixtures it would only affect a small correction. Another source of error is that it is difficult to make accurate observations on a rapidly falling mercury-in-glass thermometer. CALLENDAR advocates the use of a fairly large calorimeter, the surface of which, as well as that of the enclosure, should be permanently blackened, so as to increase the rate of loss of heat by radiation as much as possible compared with those by conduction and convection, which are less regular. For accurate work the liquid should be stirred continuously, and the outer vessel covered with a hollow lid containing water maintained at a constant temperature.

Academic Problems on the Method of Cooling.—In these it is often assumed that the specific heat of the substances used and of the material of the calorimeter are each constant. Under these conditions the equations for the heat loss may be integrated and the specific heat of a liquid found by comparing the times required for the liquid and the water to fall through the same range of temperature.

Thus
$$(m_1 s_1 + \mathrm{MS})\frac{d\theta}{dt} = af(\theta)$$
gives
$$(m_1 s_1 + \mathrm{MS})\int_{\theta_1}^{\theta_2}\frac{d\theta}{f(\theta)} = a\Big[t\Big]_1^2 = a(t_2 - t_1).$$
Similarly, for another liquid, cooling from θ_2 to θ_1, in time $t_4 - t_3$
$$(m_2 s_2 + \mathrm{MS})\int_{\theta_1}^{\theta_2}\frac{d\theta}{f(\theta)} = a(t_4 - t_3).$$
Hence
$$\frac{m_1 s_1 + \mathrm{MS}}{m_2 s_2 + \mathrm{MS}} = \frac{t_2 - t_1}{t_4 - t_3}.$$

Ferguson and Miller's Method for the Determination of the Specific Heat of Liquids.—The apparatus to be described was designed so that any reasonably equipped works laboratory could obtain values for the specific heats of liquids which could be depended upon to 1 per cent. or better. Moreover, it can be used in precision work. In brief, the method consists in supplying energy, by means of a coil heated electrically, at a rate sufficient to maintain a calorimeter and its contents at steady temperatures, for example, 5° C., 10° C., 15° C., . . . higher than the walls of an enclosure in which it is supported—but from which it is insulated thermally. After the temperature of the calorimeter and its contents has been raised to a temperature of about 50° C., the heating is discontinued, and a cooling curve obtained. If V volt. is the potential difference across the heating coil when the current through it is I amp., then $-\frac{dQ}{dt}$, which is the rate of loss of heat from the calorimeter, and which must be equal to the rate at which energy is supplied when the temperature is constant, is given by
$$-\frac{dQ}{dt} = -(\mathrm{M}s + \mathrm{C})\frac{d\theta}{dt} = \frac{\mathrm{VI}}{\mathrm{J}}, \qquad \text{(i)}$$
where M is the mass of liquid, s its specific heat, C the thermal capacity

of the calorimeter, θ the temperature difference between the calorimeter and its surroundings, and J is the mechanical equivalent of heat. When $-\frac{d\theta}{dt}$ has been determined, we have an equation from which the specific heat of the liquid used at a temperature θ above that of its surroundings can be derived.

To determine $\frac{d\theta}{dt}$ accurately from the cooling curve the method of drawing tangents, as explained on p. 213, is not good enough. Ferguson and Miller therefore assumed that $-\frac{d\theta}{dt} = k\theta^{\frac{5}{4}}$, where k is a constant. Thus (1) becomes

$$(\mathrm{MS} + \mathrm{C})k\theta^{\frac{5}{4}} = \frac{\mathrm{VI}}{\mathrm{J}} \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

To establish the validity of the equation assumed for the rate of fall in temperature, we have, by integration of that equation,

$$\theta^{-\frac{1}{4}} - \theta_0^{-\frac{1}{4}} = \tfrac{1}{4}kt, \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

where θ_0 is the initial temperature excess corresponding to the zero value of t. Now the experiment has furnished us with a set of values of θ and t: hence if $\theta^{-\frac{1}{4}}$ is plotted against t, a straight line should be obtained if the law assumed for the rate of fall in temperature is correct. Such a line was found, and from its slope the value for k determined.

The thermal capacity of the calorimeter was determined experimentally by filling the calorimeter with water and carrying out an experiment on the lines indicated.

The advantages claimed for this method are that it is easy in application, is fairly rapid, does not demand any excessive quantity of liquid, and, with simple apparatus of the ordinary laboratory type, gives results in which the error is less than one per cent. Also if the thermal capacity of the calorimeter is determined independently, the method becomes a laboratory one for the determination of Joule's equivalent.

The Specific Heats of Gases.—When heat [thermal energy] is imparted to a gas the resulting change in temperature depends upon the manner in which the gas is permitted to expand, for during an expansion the gas will do work against the external pressure the amount of which may be a very considerable fraction of the whole energy imparted to the gas. Hence, if we are to define the specific heat of a gas the conditions under which the heating takes place must be stated, for to each possible mode of expansion there is a corresponding specific heat of the gas. The two principal specific heats of a gas are the ***specific heat at constant volume*** (c_v) and that at ***constant pressure*** (c_p). If m is the mass of a gas, θ the rise in temperature when it is heated under the condition that its volume remains constant, the heat required is $mc_v\theta$, where c_v is the ***specific heat of the gas at constant volume***. Similarly, if c_p is the specific heat of the gas at constant pressure, $mc_p\theta$ is the heat required to raise by θ the temperature of a mass m of the gas when the pressure remains constant.

To show the significance of the difference between these two specific heats, consider one gram of gas contained in a cylinder fitted with a frictionless piston. If A is the area of this piston and p the external pressure, the force acting upon it is pA. If a quantity of heat is supplied sufficient to raise the temperature 1° C. under the condition that the volume of the gas remains constant, this quantity is numerically equal to the specific heat of the gas at constant volume. All the heat imparted is utilized in increasing the thermal energy of the molecules and no work is done by the gas against the external pressure, since the piston does not move.

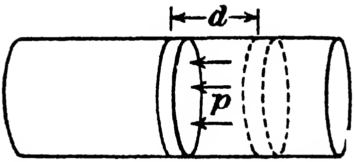


Fig. 10·5.

On the other hand, when heat is supplied to the one gram of gas under the condition that the pressure remains constant the piston will move forward a distance d, Fig. 10·5, the volume changing from v_1 to v_2. The work done against the external pressure is $pA \times d = p(v_2 - v_1)$. The heat which has gone to increase the kinetic energy of the molecules is therefore $c_p - ap(v_2 - v_1)$, where a is a conversion factor—it is equal to the reciprocal of J the mechanical equivalent of heat [cf. p. 272]. If the energy is a function of temperature only, the above increase in energy must be c_v, for c_v cal. at constant volume increase the molecular energy by the same amount. Hence it follows that the specific heat at constant pressure (c_p) is greater than the specific heat at constant volume (c_v).

Regnault's Method for Determining the Specific Heat of a Gas at Constant Pressure.—The apparatus used by Regnault is shown in Fig. 10·6. The gas to be investigated was compressed in a reservoir, A, placed in a large tank of water so that its temperature could be kept constant. The pressure in the reservoir was indicated by a mercury manometer, B. Some preliminary experiments were conducted to find what mass of gas had flowed from the container when the pressure fell between two definite limits. The flow of gas was controlled by a valve, C, and the gas then passed through a long copper spiral immersed in a thermostat, D, where its temperature was raised to that of the thermostat, a fact which Regnault established in a series of subsidiary experiments by placing a thermometer in the tube and comparing its reading with that of the thermometer in the oil. The flow of gas was kept steady in this part of the apparatus by controlling C. The pressure was indicated by the mercury manometer, G. After leaving the heater the gas passed into a thin brass vessel placed in a water calorimeter, E. Actually, this vessel consisted of a number of chambers with partitions to increase

its surface area so that there might be a rapid transfer of heat from the gas to the water. Finally, the gas escaped through a spiral tube, F, into the atmosphere. Let θ_1 and θ_2 be the initial and final temperatures of the calorimeter and its contents. If t is the temperature of the incoming gas the first portion of the gas was cooled from t to θ_1 while the last portion cooled from t to θ_2. The average fall in temperature was therefore $[t - \frac{1}{2}(\theta_1 + \theta_2)]$. The heat lost by the gas was therefore $mc_p[t - \frac{1}{2}(\theta_1 + \theta_2)]$, where m is the mass of gas passing through the calorimeter during an experiment. The calorimeter was protected from heat radiated from D by the screen shown. The experiment lasted for some time so that the errors due to heat lost

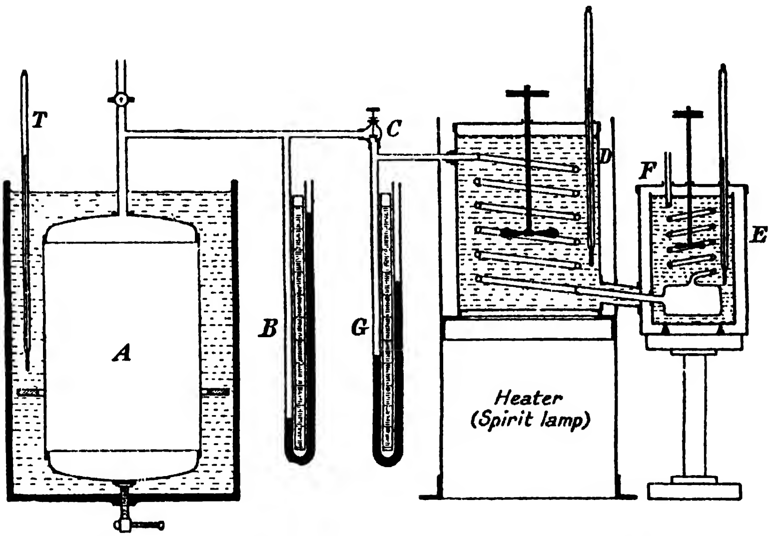


FIG. 10·6.—Regnault's Apparatus for Determining c_p.

by the calorimeter were considerable, and although a correction was made, the results obtained by Regnault differ by 2 per cent. from more recent values.

If C is the thermal capacity of the calorimeter and its contents, the heat imparted to it is $C(\theta_2 - \theta_1)$. Hence [cf. p. 288] c_p may be calculated from the equation

$$mc_p[t - \tfrac{1}{2}(\theta_1 + \theta_2)] = C(\theta_2 - \theta_1).$$

Callendar's Method for the Determination of c_p.—CALLENDAR, and more recently SWANN, utilized the method of continuous flow calorimetry to determine the specific heat of a gas at constant pressure. The essential features of their apparatus are shown in Fig. 10.7. B and C are two glass tubes about 2 cm. in diameter

joined together by a short spiral of glass tubing and surrounded by a wider tube D, the space between D and the inner tubes being exhausted to diminish the heat lost from the calorimeter. A steady stream of gas, from a thermostat, entering at E and escaping at F was heated by an electric current passing through the heater G. To measure the energy [heat] dissipated in G, an ammeter A, and resistance R, were placed in series with G, while a voltmeter V measured the potential difference between the ends of the heating coil. In their actual research these were measured by a potentiometer method so that V and A must be regarded as a symbolic representation of the apparatus actually used. The temperature of the incoming gas was measured by a platinum thermometer P_1 and after the heated gas had been 'mixed' by the wire gauze H, in order that its temperature might be uniform, a second platinum thermometer P_2 measured this temperature. We have already seen that the energy dissipated per second by the current in G is VI joule.

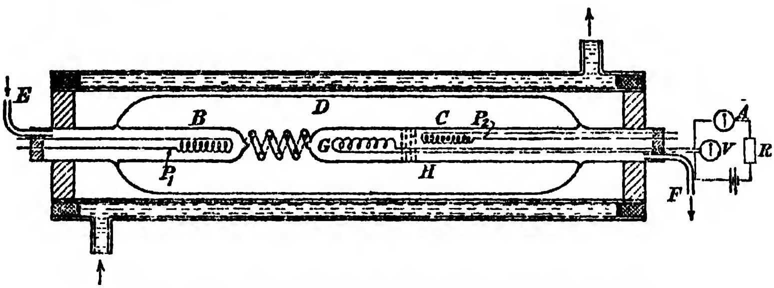


FIG. 10·7.—Callendar's Apparatus for determining c_p.

or $\frac{VI}{J}$ cal., where J is the mechanical equivalent of heat in joule. cal.$^{-1}$ If m is the mass of gas passing per second, θ the rise in temperature, we have

$$VI = J[mc_p\theta + h\theta]$$

where h is the heat loss per second per unit rise in temperature. If, therefore, the experiment is repeated so that the mass of gas flowing per second is different but that the rise in temperature is the same (by adjusting I) we have, if the suffixes refer to the two separate experiments,

$$V_1I_1 = J[m_1c_p\theta + h\theta],$$

and

$$V_2I_2 = J[m_2c_p\theta + h\theta],$$

from which c_p may be determined. Special methods were adopted to measure m_1 and m_2: we cannot discuss them here.

In order to make the conditions of the experiment quite definite the whole apparatus was surrounded by a water jacket kept at the temperature of the incoming gas.

The Specific Heat of Superheated Steam.—The specific heat of steam at constant pressure over a range of temperature from 104° C. to 115° C. was measured by Brinkworth, who used an apparatus similar to that just described but specially adapted to suit the particular purpose in view. Only one platinum thermometer was used—first to measure the temperature of the steam before any energy was dissipated in the heating coil: secondly to measure the temperature of the steam when energy was dissipated at a known rate. The steam was then condensed so that the mass of steam flowing per second was easily obtainable. The value of c_p finally obtained was 0·487 cal. gm.$^{-1}$ deg.$^{-1}$ C.

Latent Heat.—When the temperature of a solid is gradually raised, the stage at which liquefaction takes place is quite definite for all pure crystalline substances, e.g. ice, tin, but not for substances like wax, butter, some metallic alloys, etc., which gradually become plastic in character. The definite temperature at which liquefaction takes place is called the melting-point—or fusing-point. During the process of fusion a definite quantity of heat [known as the latent heat of fusion] is absorbed per unit mass of substance, and this heat is emitted in equal amount during the reverse process. The heat emitted when a mass m of liquid is caused to solidify, without change of temperature, is equal to ml, where l is termed *the latent heat of the liquid at its freezing-point.* The dimensions of l are given by

$$[l] = [\text{cal. gm.}^{-1}],$$

if a calorie is the unit of heat and the mass is expressed in grammes. It is not correct to speak of the latent heat of ice [or indeed any solid]—it is the water at 0° C. which possesses the latent heat—not the ice. Similar remarks can be made concerning the transition of liquid into a vapour at the same temperature—in this case we have the latent heat of the vapour. If m is the mass of vapour condensing to form a liquid at the temperature of the vapour, then the heat given out is ml, where l is termed *the latent heat of the vapour.* Again, it is not correct to speak of the latent heat of a liquid at its boiling-point—it is the vapour at that temperature which possesses the latent heat.

Spirits, such as Eau de Cologne, are used to alleviate headache, on account of the cooling which takes place when they evaporate. Eau de Cologne is used because it evaporates easily [i.e. it has a high vapour pressure, cf. p. 245] and its latent heat is considerable. Menthol is sold for the same purpose.

Experimental Determination of the Latent Heat of Fusion of Ice, or the Latent Heat of Water at 0° C.—The mass of a calorimeter is found, and also the mass of water required to fill it

to the extent of two-thirds its volume. The calorimeter is placed on corks in a glass vessel and the temperature observed. One or two lumps of ice, previously dried, are then added ; the temperature is noted when all the ice has melted, and stirring produces no further effect on the thermometer. The mass of the calorimeter and its contents is again determined, in order to ascertain the quantity of

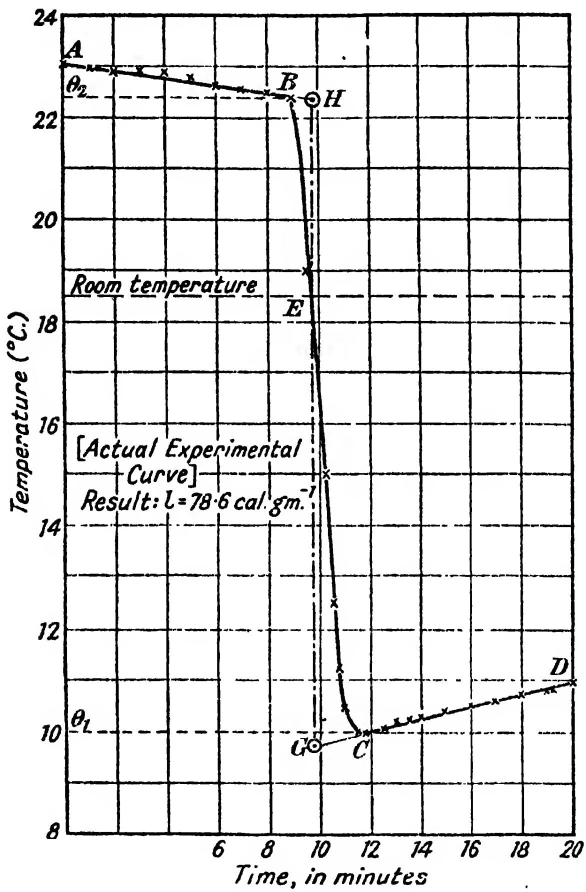


FIG. 10·8.

ice used. Approximate results having been thus obtained, the initial temperature of the calorimeter should be adjusted, so that the final temperature is likely to be a few degrees below room temperature so that Ferry's method of applying a correction for the heat exchange between the calorimeter and its surroundings may be used. If the temperature becomes much less than that of the

room there is a tendency for dew to be deposited on the calorimeter, and an unknown amount of heat imparted to it.

Let AB, Fig. 10·8, represent the cooling curve for the calorimeter before the ice is added. When the ice is introduced, at the stage represented by B, the cooling becomes more rapid and the temperature falls until all the ice has melted and the temperature corresponding to the point C reached. The temperature then rises as indicated by the portion CD of the curve. Let E be the point of intersection of the 'room-temperature' line with ABCD, and suppose that AB and DC produced cut the vertical line through E in H and G respectively. Then HG is the fall in temperature, θ, (say), which would have been obtained in the absence of any heat exchange between the calorimeter and its surroundings.

While the ice is melting it should be kept below the water surface by means of a stirrer to which a piece of copper gauze has been fixed. This also acts as an efficient stirrer throughout the experiment.

If C is the thermal capacity of the calorimeter and its contents, the heat abstracted from them is $C\theta$. This is used in melting a mass, m, of ice—found by weighing the calorimeter at the conclusion of the experiment—and in raising its temperature to θ_1, i.e. to C. Thus,

$$C\theta = ml + (m \times 1 \times \theta_1) = m(l + \theta_1),$$

so that l may be found.

The chief objection to this method of determining the latent heat of fusion of ice is that it is never certain whether or not free water is associated with the ice which is introduced into the calorimeter. This difficulty has been avoided as follows:—Some broken pieces of ice were cooled several degrees below zero and their mass (about 100 gm.) determined. It was then introduced into a calorimeter containing kerosene oil at a temperature two or three degrees below 0° C. Under these circumstances the ice must have been entirely free from water. A very small electric current was passed through a heating coil immersed in the oil to raise the temperature of the calorimeter and its contents to $-1°$ C. Electrical energy was then supplied at a much greater rate and sufficient to melt the ice and raise the temperature of the calorimeter and its contents to 0·5° C. In this way the ice had certainly all been melted and the heat dissipated in the calorimeter had been employed in three ways:

(i) in raising the temperature of the calorimeter and its contents from $-1°$ C. to 0° C. [specific heat of ice = 0·493 cal. gm.$^{-1}$ deg.$^{-1}$ C.]

(ii) in melting the ice, and

(iii) in raising the temperature of the water, oil, and calorimeter to 0·5° C.

In addition, there was a small exchange of heat between the calorimeter and its surroundings. The sign of this will depend upon circumstances, but a correction for this was made. To make this correction as small as possible, the calorimeter was placed in a well-lagged box maintained at 0° C.

In addition to the advantage which this method possesses over the one described above, it also has the following—the fluidity of the oil remains practically unchanged, and there is no thermal reaction between the oil and ice or water.

The above experimental procedure is due to A. W. Smith, who obtained 78·896 mean cal. gm.$^{-1}$ as the latent heat of fusion of ice. The value generally accepted to-day is 79·70 cal. gm.$^{-1}$

The Latent Heat of a Vapour.—The latent heat of the vapour of water at its normal boiling-point can be investigated experimentally by means of the apparatus shown in Fig. 10·9. The underlying idea is to pass the vapour [steam] into a known mass of liquid [water] contained in a calorimeter. From the observed rise in temperature of the calorimeter and its contents, the latent heat can be calculated.

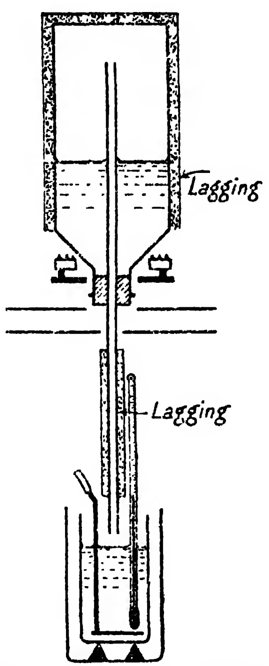


Fig. 10·9.—Latent Heat of a Vapour.

Into a copper vessel with a narrow neck there is fitted a glass tube [preferably made of 'pyrex,' because it withstands changes of temperature much better than ordinary glass]. The vessel is half filled with the liquid under investigation and then placed upon a ring burner standing upon a sheet of asbestos. A quantity of the liquid cooled to a temperature several degrees below that of the room, but not sufficiently cold for moisture from the atmosphere to become deposited on the calorimeter, is placed in a calorimeter and its mass found. It is desirable that the calorimeter should be about two-thirds filled. When the vapour is issuing freely from the tube of the boiler, the temperature of the calorimeter is observed, and the vapour allowed to play directly on the surface of the liquid where some of it is condensed. The steam that escapes does not impart heat to the calorimeter, and so does not concern us. By obtaining a

time-temperature curve for the calorimeter before the passage of the steam, during its passage, and afterwards, the rise in temperature in the absence of any heat exchange between the calorimeter and its surroundings may be found [cf. p. 207].

If C is the thermal capacity of the calorimeter and its contents, θ_1 and θ_2 its initial and final temperatures, θ the corrected rise in temperature, t the steam temperature, then the heat given out by a mass m of steam in condensing to water at temperature t and then cooling to θ_2, is $ml + (m \times 1 \times \theta_2)$. But this is $C\theta$, so that l may be found.

On Laboratory Methods for the Determination of Latent Heats.—Instead of using a copper calorimeter and somewhat small quantities of water a Dewar flask and larger quantities of water may be used. It is first necessary to determine the thermal capacity of the flask. To do this a known mass of water at 25° C. is placed in the flask and, after thoroughly shaking, its temperature is noted. An approximately equal amount of water of known mass and at a temperature of about 10° C. is then added to the flask and after it has been thoroughly mixed with the other water the final temperature is recorded. The heat exchange between the flask and its surroundings is small and may be neglected. The thermal capacity of the flask is then calculated.

To determine the latent heat of fusion of ice a known mass of water at 25° C. is placed in the flask, well shaken, and the temperature noted. Pieces of dry ice are then added until the temperature is reduced to about 10° C. The mass of ice used is then determined and the latent heat of fusion calculated in the usual way.

The latent heat of steam at the boiling-point of water under existing atmospheric conditions of pressure may be determined in a like manner.

The masses of ice or of steam used in these experiments is considerably greater than if an ordinary calorimeter is used, and although the changes in temperature are rather large the heat exchange between the calorimeter and its surroundings is small and negligible.

Berthelot's Apparatus.—The liquid whose heat of vaporization is required is placed in a special flask, A, Fig. 10·10. A glass tube open at both ends projects through the bottom of this flask. A ground glass joint, C, connects this tube to a glass spiral and bulb immersed in water in a calorimeter. The liquid is heated by a small gas-ring, G, the direct passage of heat from the flame to the calorimeter being prevented by a wooden cover held in position by an outer vessel which diminishes the loss of heat from the calorimeter by convection and conduction. When the liquid boils the vapour condenses in the spiral and is collected in the bulb, D. When a rise in temperature of about 5° C. has been obtained the flame is put out and the final temperature noted. S is a stirrer and T a mercury thermometer reading directly to 0·1° C. The net exchange of heat between the calorimeter and its surroundings may be considerably reduced by commencing the experiment with the temperature say three degrees

below that of the room and passing the vapour until the temperature is three degrees above—there is no need to make a preliminary experiment in this instance. The bulb is removed, the two exits being closed with corks to prevent evaporation. The mass of vapour condensed is found and its latent heat calculated as follows.

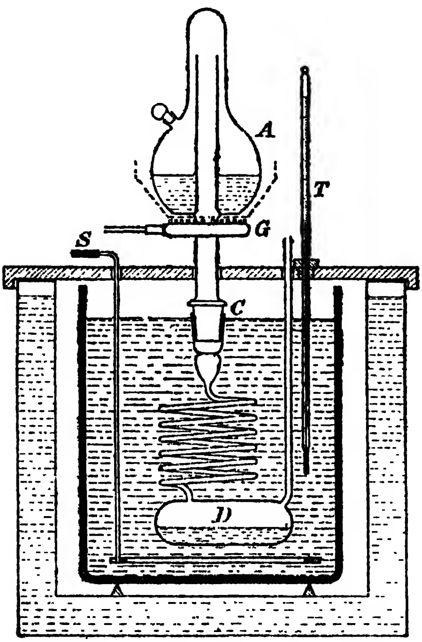


FIG. 10·10.—Latent Heat of a Vapour (Berthelot's apparatus).

Let m be the mass of liquid, L its heat of vaporization, s the specific heat of the liquid [assumed constant over the temperature range involved], T the boiling point of the liquid under the pressure conditions prevailing, while θ_1 and θ_2 are the initial and final temperatures of the calorimeter. Then the heat given out by the vapour + the heat given out by the liquid produced in cooling to θ_2 is $ml + ms(T - \theta_2)$. The heat absorbed by the calorimeter and its contents, of thermal capacity C, is $C(\theta_2 - \theta_1)$. Hence

$$ml + ms(T - \theta_2) = C(\theta_2 - \theta_1)$$

l may be calculated when the other factors in this equation are known.

An Electrical Method for determining the Latent Heat of a Vapour produced at Atmospheric Pressure.—The apparatus consists of an inverted Dewar flask, F, Fig. 10·11 (a), provided with a rubber bung through which pass a delivery tube G, surrounded by a Liebig's condenser, and a tube H of the form indicated. A coil of manganin wire is placed round the lower portion of that part of G which is inside the flask and connected to a battery as shown. The current I amp. through the wire is measured by the ammeter A, while a voltmeter V gives the potential difference across the wire. The liquid under investigation is introduced into the flask and boiled. Care must be taken to see that the liquid level in H is always above the top of the manganin coil—otherwise the vapour will be superheated. The mass of liquid evaporating

in 10 minutes is determined: let m be the mass evaporating per second. Then if l is the latent heat of vaporization

$$\frac{\mathrm{VI}}{\mathrm{J}} = (ml + h),$$

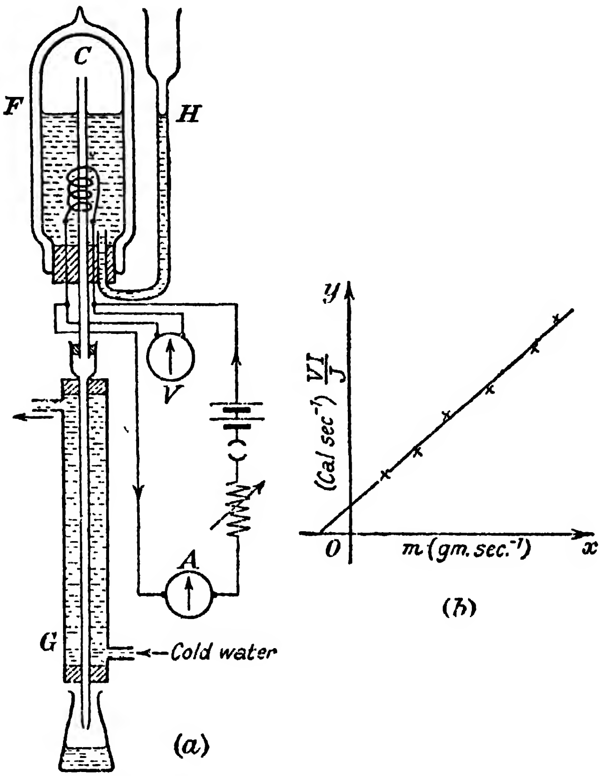


FIG. 10·11.—Laboratory Method (Electrical) for determining the Latent Heat of Vaporization of a Liquid.

where J is the mechanical equivalent for heat (4·18 joule. cal.$^{-1}$) and h is the heat lost per second. If therefore a series of observations is made and a graph showing the relation between m and $\frac{\mathrm{VI}}{\mathrm{J}}$ drawn, the slope of the line will be l. An actual graph obtained in the laboratory is shown in Fig. 10·12 (b). It gave $l = 213$ cal. gm.$^{-1}$ for alcohol.

Continuous Flow Method of Determining Latent Heats of Vapours.—A modification, suitable for students' use, of a more recent form of apparatus for determining the latent heats of vaporization of liquids at their normal boiling-points is shown in Fig. 10·12. It consists of a well-lagged heating vessel, A, in which the liquid is placed.

The thermal energy is supplied by means of an electric current passing through a coil of resistance wire which surrounds the lower part of A.

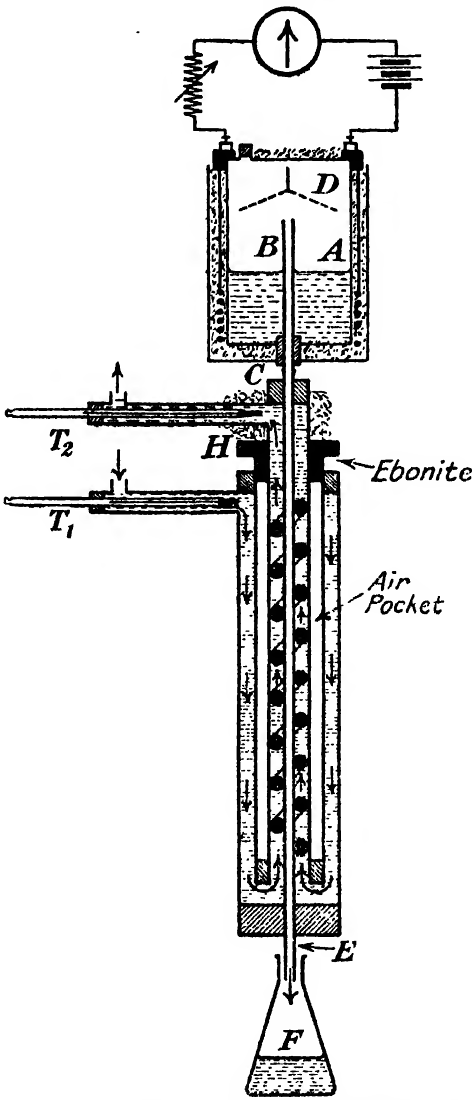


FIG. 10·12.—Modern Apparatus (Steady Flow Method) for Determining Latent Heat of Vaporization of a Liquid.

The coil is insulated electrically from A by means of asbestos paper. The terminals at the ends of the copper leads to the heating element are fixed in an ebonite ring attached to A. A glass tube, B, passes

through the base of A, and it is of sufficient length to project beyond the surface of the liquid. A conical shield, D, placed inside the heating vessel diminishes any transfer of radiant energy between the liquid and the top of the boiler; perforations in this metal shield permit the vapour to circulate freely. B is connected by means of a conical joint, C, to a long glass tube E surrounded by a condenser through which a stream of cold water is passed. This condenser has the particular form indicated so that any heat lost from the water after it has been heated is given to the incoming water: in this way heat exchanges between the apparatus and its surroundings are minimized. [The ebonite, or hard wood, collar H permits the apparatus to be assembled quite easily.] Let θ_1 and θ_2 be the initial and final temperatures of the water as measured by the thermometers T_1 and T_2. Then $M(\theta_2 - \theta_1)$ is the heat lost per second by the vapour and the liquid it forms, if M is the mass of water flowing per second. If l is the latent heat of vaporization of the liquid, and s its mean specific heat over the range of temperature from θ_1 to θ_3 where θ_3 is the boiling-point of the liquid, then the above quantity of heat is also

$$m[l + s(\theta_3 - \theta_1)],$$

where m is the mass of liquid condensing per second. This is found by weighing the amount of liquid which collects in the conical flask F in a known time. Hence

$$M(\theta_2 - \theta_1) = m[l + s(\theta_3 - \theta_1)],$$

so that l may be calculated.

This equation is not exact, since the temperature of the liquid as it leaves the tube E is not θ_1; however, the correction is small.

Since two different thermometers are used to measure θ_1 and θ_2, a correction is necessary for the fact that no two mercury thermometers are consistent in their indications. To estimate the correction to be applied, water at temperature θ_1 is passed through the apparatus while no vapour is condensing. If T_2 reads ϕ_1, the temperature difference to be used in the above equation is $(\theta_2 - \phi_1)$. The correction, on this account, to $\theta_3 - \theta_1$ is negligible.

The Ice Calorimeter.—The first and simplest form of ice calorimeter was invented by BLACK. It consisted of a block of ice, free from air bubbles, in which a cavity had been made. The top of the calorimeter was covered with a slab of ice so that a body placed inside the calorimeter was thermally insulated from external objects. A known mass of solid whose specific heat was required was heated to some definite temperature and then transferred to the cavity which had been dried by wiping it with some absorbent. The temperature of the solid was soon reduced to 0° C., a definite amount of ice being melted in the process. The amount of water formed was estimated by wiping the cavity with a sponge—the increase in mass, say M, was then found. If l is the latent heat of fusion of ice, the heat given to the ice is Ml. This is equal to $ms\theta$ where m is mass of the solid and θ the initial temperature of the heated body. Whence

$$s = Ml \div m\theta.$$

The objections against this simple apparatus are that it is difficult to remove all the water melted and also not easy to obtain large blocks of homogeneous ice. LAVOISIER and LAPLACE improved the ice calorimeter, but its value was enhanced when BUNSEN devised the form shown in Fig. 10·13. A tube P is fixed into the upper end of a wider tube Q, shaped as shown. The space between them is filled with air-free water, with the exception of the lower portion and the side tube S which contain mercury. A capillary tube R, graduated so that the volume between any two marks upon

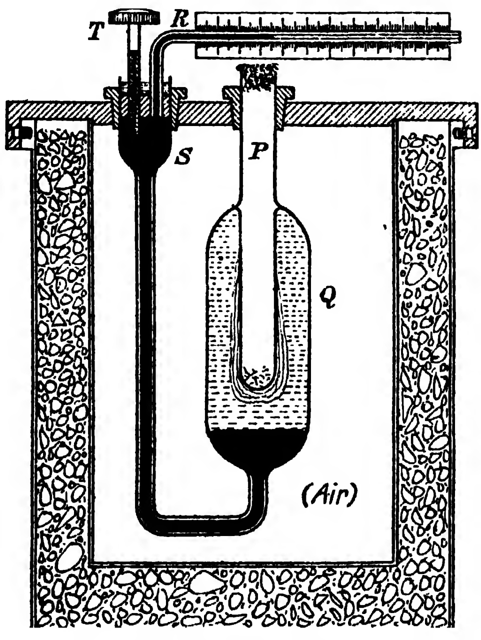


FIG. 10·13.—Bunsen's Ice Calorimeter.

it is known, is inserted through an ebonite stopper at S, the end of the mercury thread in R being adjusted to a convenient position by means of the iron screw T. The whole apparatus, with the exception of R, is placed in a large Dewar vessel containing melting ice and left overnight so that its temperature is everywhere 0° C. A little ether is then placed in P and caused to evaporate quickly by bubbling air through it. The heat required for this evaporation is partly abstracted from the water at 0° C. so that some ice is formed around P. When sufficient ice has been formed the apparatus is placed inside the double-walled container shown. The space between the above walls is filled with melting ice so that eventually

the temperature is everywhere 0° C. A small quantity of water at 0° C. is then placed in P and a hot body of mass m at temperature θ lowered into P—a swab of cotton-wool at the bottom of P diminishes the risk of fracturing the apparatus. During the cooling of the hot body a quantity of heat $ms\theta$ is emitted, thereby melting some ice. The volume of the water in Q is thus altered, the change in volume being derived from the observed change in the position of the mercury in R. Let d be the distance through which the mercury recedes, and a the cross-sectional area of the capillary tube. Then the diminution in volume is (ad). Let ρ_1 be the density of water at 0° C., ρ_2 that of ice at the same temperature. The specific volumes of water and ice at 0° C. are therefore $\frac{1}{\rho_1}$ and $\frac{1}{\rho_2}$, respectively. The contraction when 1 gm. of ice melts is therefore $\frac{1}{\rho_2} - \frac{1}{\rho_1}$. Hence the mass of ice melted is $ad \div \left[\frac{1}{\rho_2} - \frac{1}{\rho_1}\right]$ gm., for which the heat required is $\frac{lad\rho_1\rho_2}{\rho_1 - \rho_2}$ cal.

$$\therefore\ ms\theta = \frac{lad\rho_1\rho_2}{\rho_1 - \rho_2},$$

or

$$s = \frac{lad\rho_1\rho_2}{(\rho_1 - \rho_2)m\theta}.$$

For this instrument to be used successfully it is necessary to have an air space between the actual calorimeter and the ice in the double-walled container—heat then only passes slowly between the calorimeter and the container, so that the creeping of the mercury along R which occurs if this precaution is not adopted is much reduced. CALLENDAR surrounded Q by a second glass bulb, the intervening space being exhausted. The rate at which the creeping occurred was thereby still further diminished.

The real source of trouble in using Bunsen's ice calorimeter lies in the fact that the ice inside the calorimeter melts at a temperature below 0° C. because of the pressure exerted by the mercury column (due to its weight and surface tension).

The Bunsen ice calorimeter may be standardized by first carrying out an experiment with water. For this purpose a known mass of water at, say, 20° C. is introduced into the calorimeter, and this water, in cooling down to 0° C., supplies a definite amount of heat to the calorimeter. The distance through which the mercury in the capillary recedes is noted and the so-called constant for the calorimeter calculated. This constant is defined as the heat required to be given to the calorimeter in order that the mercury shall recede 1 mm. in the capillary tube.

The process of standardization as just described is difficult. It

is much better to use an electrical method. To do this a coil of manganin wire, wound on a mica frame, is introduced into the bulb of the calorimeter, containing a thin insulating oil, and a steady electric current of I amperes passed through it for t seconds. Let V volt. be the difference in potential across the coil as measured with a voltmeter. Then the heat liberated will be $\frac{VIt}{J}$ calories, where J is the mechanical equivalent of heat [cf. p. 272]. If the corresponding change in the position of the mercury is observed, the calorimeter constant may be deduced at once. The advantage of standardizing the instrument by either of the above methods is that it is not necessary to assume a value for the density of ice.

Joly's Steam Calorimeter.—This calorimeter was designed by Joly in 1886 as an accurate means of determining the specific heat of a solid. The essential parts of this calorimeter are shown in Fig. 10·14. A metal enclosure, A, called the steam chamber, is furnished with a wide side tube, B, through which steam is passed, and

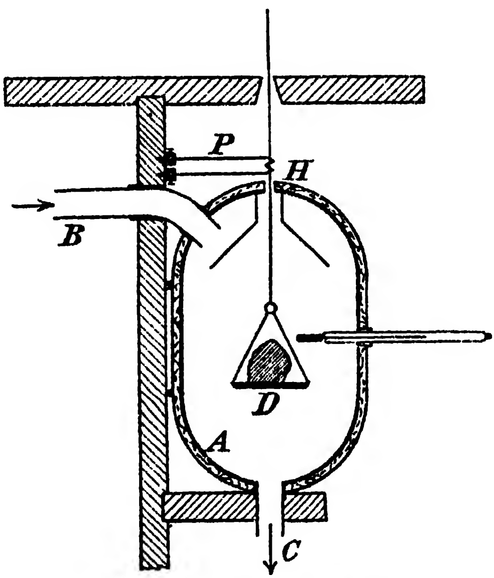


FIG. 10·14.—Joly's Steam Calorimeter.

an exit tube, C, near the bottom of the chamber. A small metal pan, D, is suspended from one arm of a balance by means of a fine wire, passing through a small hole in the top of the chamber. The solid, whose specific heat is required, is placed on this pan and its mass determined; the temperature is recorded. Steam is now admitted; some condenses on the pan, solid, and sides of the

chamber. The increase in mass of the pan and solid is due to the moisture which has condensed upon them in raising the temperature to that of the steam. This increase is determined about five minutes after the entry of the steam—if the experiment is allowed to continue, drops of water from the top of the chamber fall upon the pan and, by so doing, vitiate the results. The hole must not be too large at the point H where the supporting wire passes into the steam chamber, for then moisture may condense on the balance, and if it is small, steam condenses there so that the balance readings are not steady. The condensation of the steam is prevented by placing a nickel coil, P, round the wire, immediately above this hole. The coil is heated to redness by an electric current, thus preventing the formation of a water globule.

If μ is the mass of steam condensed, l its latent heat, the quantity of heat imparted to the pan and solid is μl. This is also the heat required to raise the pan from its initial temperature θ_1 to its final temperature θ_2, together with the heat required to raise the solid through the same range of temperature. Hence

$$\mu l = (\mathrm{MS} + ms)(\theta_2 - \theta_1),$$

where M and m are the masses of the solid and pan of specific heats S and s respectively. In general, the pan is made of copper so that its specific heat is known; should it be unknown, however, it can be determined by a preliminary experiment in which no solid is placed in the pan.

In practice a small correction to the above equation has to be made owing to the fact that the body is first weighed in air and then in steam.

The Specific Heat of a Gas at Constant Volume.—For this determination it is essential to enclose the gas in a container the mass of which is, in general, much greater than that of the enclosed gas. Joly's differential steam calorimeter, Fig. 10·15, was the first means whereby the specific heat c_v of the gas could be measured directly and not deduced from the difference of two thermal capacities, viz. that of the gas and its container, and that of the container alone. The apparatus consisted of two equal copper spheres, SS, 6·7 cm. in diameter, suspended in the same steam chamber from the opposite arms of a balance. These spheres were provided with pans, AA, to trap any condensed steam which might fall from the spheres. One of the spheres contained gas at normal pressure while the second contained some of the same gas at a pressure of several atmospheres [pressures of 20 atmospheres were sometimes used]. The initial temperature having been recorded, steam was passed into the chamber through D and the mass necessary to restore equilibrium was due to the excess condensation brought about by the difference

between the mass, m, of gas in the two spheres. If l is the latent heat of steam, and μ the difference in the mass of steam condensed on the two spheres, the heat given to the m gm. of gas is μl. This is equal to $mc_v(\theta_2 - \theta_1)$, where θ_2 is the temperature of the steam and θ_1 that of the calorimeter initially.

Since the spheres are equal the buoyancy correction to which reference has been made [cf. p. 231] is eliminated. If results of high

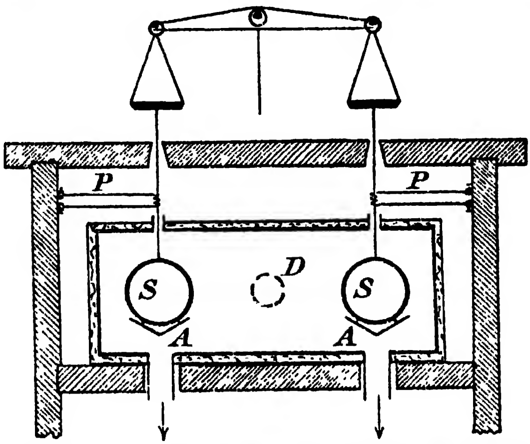


FIG. 10·15.—Joly's Differential Steam Calorimeter.

precision are being aimed at, a small correction has to be made for the expansion of the copper spheres. Also, if the two spheres are not exactly equal an experiment must first be made with each sphere exhausted in order to ascertain the difference in the amount of steam condensed when the temperature of the spheres is raised to the steam temperature. A correction for this difference is then applied when the main experiment is carried out.

Further Remarks about the Steam Calorimeter.—One advantage of the steam calorimeter is that it may be used to determine the specific heats of solids, liquids, and gases; moreover, it applies whether or not the material is available in large or small quantities. Liquids, powders, and substances attacked by water vapour, must be sealed in a glass container, the thermal capacity of which must be known—it may be determined by this method of calorimetry. The method is not only an accurate one, but it is also universal in its applications.

Atomic Heats.—In 1818 DULONG and PETIT enunciated the law which bears their names, viz. the product of the specific heat of an element in the solid state and its atomic weight is constant and equal to 6·4. If a quantity of material equal to its atomic weight

in grams, i.e. one gram-atom, is considered the above law implies that the thermal capacity of every gram-atom is the same. Since every gram-atom of substance contains the same number of atoms it follows that the atoms of all elements in the solid state (whatever that means) have identical thermal capacities.

It was soon found that the law was only a first approximation to the truth, for the constant varied from 5·7 to 6·7. When we remember that the thermal capacity of a substance measures the energy [heat] necessary to cause the molecules to move more rapidly or to rotate more quickly, this result is really not very surprising.

For a long time the elements carbon, boron, and silicon were considered to be exceptions to this law, but when their specific heats were measured at high temperatures the anomaly disappeared. DEBYE has since shown that this law is really a first approximation to a law which is more complicated but at the same time more universal. A discussion of this law would take us too far from our present object if it were considered here.

Molecular Heats.—NEUMANN, REGNAULT, and others, have extended the above law so that it applied to molecules. They defined the ***molecular heat*** as the product of the specific heat and molecular weight of a substance and showed that the molecular heat of a body was equal to the sum of the atomic heats of its constituents. The molecular heats of a gas at constant pressure and at constant volume are denoted by C_p and C_v respectively.

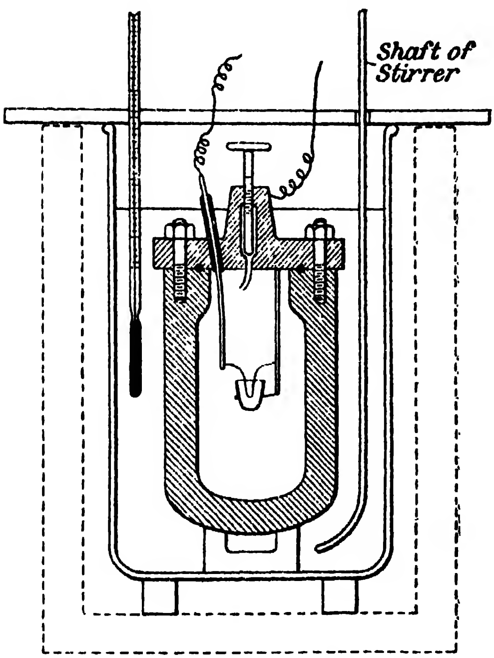


FIG. 10·16.—A Bomb Calorimeter.

Determination of the Calorific Value of a Sample of Coal.—[The calorific value of a solid or liquid fuel is the amount of heat given out per unit mass—1 lb., 1 gm. or 1 kgm.—when the substance is burnt. For a gas, it is the heat given out when a definite volume—1 standard cubic foot—is burnt.] The apparatus used for this purpose is known as a bomb calorimeter, one form of which is indicated in Fig. 10·16.

The coal is powdered, dried, and a definite mass placed in the small capsule provided. The cover is screwed in position and oxygen admitted until the pressure inside is about 20 atmospheres. The whole is placed in a metal calorimeter containing a known mass of water. This is well stirred and the temperature observed by a thermometer graduated to read to 0·01° C. The coal is ignited by passing an electric current through a platinum spiral lying in contact with the coal. The current is such that the wire is raised to incandescence. This current is only maintained for a few seconds, so that the heat developed by the current may be neglected in comparison with the other heat quantities involved. When the combustion is completed the final temperature of the water, etc., is noted and from the known thermal capacity of the calorimeter and its contents the calorific value of the sample may be derived.

Example.—A bomb calorimeter whose thermal capacity was 647 cal. deg.$^{-1}$ C. was immersed in 2000 gm. of water in a vessel whose thermal capacity was 98 cal. deg.$^{-1}$ C. After burning one gram of coal the temperature increased from 17·65° C. to 19·98° C. Calculate the calorific value of the coal.

Heat imparted to water, etc. = total thermal capacity of calorimeter and contents × rise in temperature

$$= 2745 \times 2{\cdot}33 = 6400 \text{ cal.}$$

Hence its calorific value is 6400 cal. gm.$^{-1}$

EXAMPLES X

1.—How much heat is lost by a copper block 8·1 cm. × 3·6 cm. diameter in cooling from 93° C. to − 8° C.? Specific heat of copper = 0·092 cal. gm.$^{-1}$ deg.$^{-1}$ C.; density 8·8 gm. cm.$^{-3}$

2.—A calorimeter contains 70·2 gm. of water at 15·3° C. On adding 143·7 gm. of water at 36·5° C. the temperature rises to 28·7° C. What is the thermal capacity of the calorimeter?

3.—A calorimeter of thermal capacity 8·1 cal. deg.$^{-1}$ C. contains 60·3 gm. of water at 13·2° C. A solid of mass 46·3 gm. at 99·6° C. is dropped into the calorimeter. The final temperature is 18·2° C. Calculate the specific heat of the solid assuming that 5·3 cal. of heat are lost during the course of the experiment.

4.—How much ice will be melted when a piece of metal, mass 60·4 gm., specific heat 0·042 cal. gm.$^{-1}$ deg.$^{-1}$ C. at 627° K. is dropped on to ice? [l = 80 cal. gm.$^{-1}$].

5.—A piece of metal of mass 64·2 gm. at 9·7° C. is placed in the chamber of a Joly steam calorimeter (barometer 76·9 cm. of mercury), 1·52 gm. of steam condense. If l = 540 cal. gm.$^{-1}$, calculate the specific heat of the metal.

6.—The mass of mercury required to fill 10·0 cm. of the index tube in a Bunsen ice calorimeter is 3·1 gm. The thread moves 54·6 mm. when 14·6 gm. of metal at 97·2° C. are introduced into the calorimeter. The densities of ice and of water, each at 0° C., may be taken as 0·917 gm. cm.$^{-3}$ and 1·000 gm. cm.$^{-3}$, respectively. If the latent heat of water at 0° C. is 80 cal. gm.$^{-1}$, and the density of mercury 13·6 gm. cm.$^{-3}$, calculate the specific heat of the metal.

7.—Define the terms *thermal capacity* and *specific heat.* Describe a

method other than that known as the method of mixtures of determining the specific heat of glass. Given that 1 cm.^{3} of ice at 0° C. yields 0·918 cm.^{3} of water at the same temperature and that the mercury in a Bunsen's ice calorimeter recedes 5 cm. in a capillary whose cross-section is 0·01 cm.^{2} when a body of mass 10 grams and initial temperature 100° C. is placed inside the calorimeter, calculate the specific heat of the substance. [l for water at 0° C. = 80 cal. gm.$^{-1}$]

8.—Explain why the specific heat of a gas at constant pressure is not the same as the specific heat at constant volume, and state what becomes of the energy on heating the gas at constant volume. Describe and explain a method of measuring one of these specific heats for air.

9.—Describe and explain a method, other than that known as the 'method of mixtures,' of determining the specific heat of a liquid.

10.—Describe an apparatus which, in your opinion, would be suitable for measuring the calorific value of coal gas, and explain how you would use it.

11.—Define the terms : *latent heat, specific heat.* Describe Joly's steam calorimeter and explain how it may be used to determine the specific heat of a piece of india-rubber.

12.—Describe an accurate method of determining the specific heat of aniline and discuss the advantages of this liquid when used instead of water as a calorimetric liquid.

13.—Describe an accurate method of determining the latent heat of vaporization of alcohol.

14.—A gram of ice at 0° C. contracts 0·090 cm.^{3} on melting to form water at the same temperature. A piece of metal is heated to 78° C. and then carefully inserted inside a Bunsen's ice calorimeter. If the total contraction is 0·056 cm.^{3} and the mass of the metal is 8·76 gm. calculate the specific heat of the metal. What is the thermal capacity of all the metal ? What would be the value for the specific heat of the metal if the unit of temperature were 1° F. ? [l = 79 cal. gm.$^{-1}$.]

15.—A mass of 185 gm. of copper was heated in steam when the barometer read 74·6 cm. The copper was dropped carefully into 84·5 gm. of alcohol. The temperature of the alcohol and its container rose from 16·4° C. to 23·7° C. If the specific heats of copper and alcohol are 0·10 and 0·63 cal. gm.$^{-1}$ deg.$^{-1}$ C. respectively, calculate the thermal capacity of the container.

16.—Steam at a temperature of 100° C. is carried along an iron pipe 50 metres long and weighing 20 lb. a foot. When the steam first enters the pipe the temperature of the iron is 15° C. If the specific heat of iron is 0·12 cal. gm.$^{-1}$ deg.$^{-1}$ C. and the latent heat of steam 540 cal. gm.$^{-1}$, what is the minimum amount of steam condensed to water before it passes freely along the pipe ?

17.—Describe the steady flow method of measuring the specific heat of a liquid, and explain why it is especially suitable for measuring the small variations of specific heat with temperature.

18.—Equal quantities of water are placed in two similar calorimeters, except that the outer surface of one is blackened and the outer surface of the other is polished. When the blackened calorimeter is suspended in an enclosure at constant temperature, the water in it cools from 61° C. to 59° C. in x seconds, and from 41° C. to 39° C. in y seconds. When the polished calorimeter is similarly treated the water cools from 61° C. to 59° C. in z seconds. Deduce expressions for the temperature of the enclosure and for the time the polished calorimeter takes to cool from 41° C. to 39° C. Indicate any assumptions you make.

CHAPTER XI

CHANGE OF STATE

Fusion

Normal Freezing-point.—For every substance such as water, naphthalene, the elements, and eutectics [cf. p. 242] there is a definite temperature above which the substance is wholly liquid, while below that temperature it is solid. This temperature is called the normal freezing-point when the external pressure is one atmosphere. Thus the normal freezing-point of water is 0° C.—if the applied pressure is changed the freezing-point alters. Later on, it will be found that this change is very small even when the change in pressure is one atmosphere, so that the effect of changes in atmospheric pressure on the melting-point of ice is negligible in practice. Other substances such as fats, alloys in general, silica, and glass, do not have a definite melting-point but are plastic over a range of temperature.

The Laws of Fusion.—(*a*) ***For a given external pressure the temperatures at which substances belonging to the first class of substance melt are the same as that at which they solidify, i.e. during fusion or solidification the temperature is constant.***

(*b*) ***During fusion heat is absorbed (latent heat of fusion) while during solidification heat is disengaged (latent heat of the liquid at the freezing-point).***

Determination of the Melting-point of a Solid.—*Method i :* A capillary tube is made and dipped in a crucible containing a little of the molten substance. The liquid rises in the tube which may then be withdrawn. When cold the tip of the tube is heated to redness and closed to prevent the liquid from leaving the tube when in use later. The tube is attached to a thermometer by a rubber band and placed in a beaker of water (or oil) heated by a small gas flame. The open end of the capillary should be above the surface of the liquid in the beaker. The liquid is well stirred : when the substance melts the temperature is recorded. The flame is removed and the temperature at which solidification sets in noted. The mean of these temperatures is the melting-point required. To save time a preliminary experiment should be made in which the rate at which heat is imparted to the liquid is increased. When an approximate value of the melting-point has been obtained in

this way the slow rate of heating should then be adopted commencing at a temperature a few degrees below the approximate value of the melting-point.

Method ii: When a body, such as a crucible containing a pure metal, is cooling, the curve obtained by plotting the temperature against time is regular and smooth, provided that no change of state occurs. During the passage, however, from one state to another, say from liquid to solid, a certain amount of heat is almost invariably emitted—in fact it is the latent heat of the liquid metal at its freezing-point which is given out. When the conditions are such that this change of state is about to take place, the first few molecules which separate give up their latent heat. Any further loss of heat does not cool the body but causes a further separation of solid particles, and the amount of solid which separates is just sufficient to balance the heat lost; the temperature of the mass therefore remains constant. This process continues until solidification has taken place, when the temperature again falls. Such facts are utilized in the determination of the melting-points of substances.

Experiment.—Clamp a boiling tube in a vertical position and surround the tube by boiling water. Introduce naphthalene into the tube until, when melted, it is about two-thirds filled with liquid. Insert a thermometer in the liquid so that its bulb is in the centre of the liquid and fix it rigidly in this position, no stirrer being used. Remove the water, dry the outside of the tube and surround it by a large vessel to protect it from air currents. Record the temperature at intervals

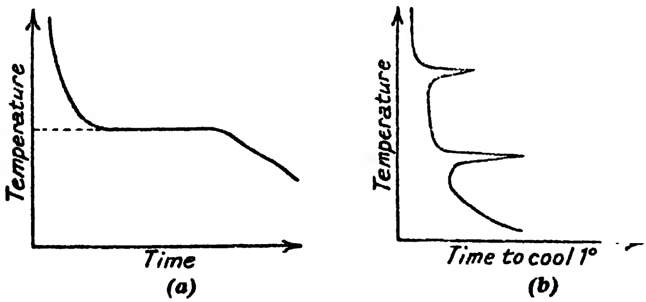


FIG. 11·1.—Ordinary Cooling Curve. Inverse-rate Cooling Curve.

of half a minute, continuing the readings until the temperature is about 20 degrees below the melting-point. When the observations are plotted, the temperature axis being vertical, the resulting curve, Fig. 11·1 (*a*), is called a *cooling curve.* If the experiment has been carried out with due care it will be noticed that one portion of the curve is parallel to the time axis. The temperature corresponding to this is the melting-point of the naphthalene.

The Inverse-rate Curve.—Metallurgists are frequently required to construct a cooling curve of an alloy because it conveys to them

information concerning its composition. Since the observations are usually made at high temperatures thermo-couples have to be used in place of mercury thermometers. Working at high temperatures means that the rate of cooling will be increased, so that if a cooling curve is constructed in the usual way it is difficult to estimate those temperatures where a change in phase takes place. To make these transitions more apparent an *inverse-rate* curve is constructed. In this the temperature is plotted vertically, but along the x-axis the time to cool one degree is plotted. The type of curve obtained is similar to that in Fig. 11·1 (*b*). Changes of state are indicated where the curve shows peaks running parallel to the x-axis.

Unstable Conditions.—Liquids with a definite freezing-point can be reduced in temperature to a point some degrees below the normal freezing-point, if the abstraction of heat takes place slowly and the liquid is not shaken. The liquid is then said to have been ***supercooled.*** Such phenomena were noticed by FAHRENHEIT as early as 1724. GAY-LUSSAC also reported that if water were placed in a clean vessel and covered with a layer of oil, the whole remained liquid at $-$ 12° C.—a slight shake and the whole froze, the temperature rising to 0° C. Pure antimony can be supercooled by 60° C. The phenomenon is easily observed in the case of sodium thiosulphate which melts in its own water of crystallization. The solution may be cooled to room temperatures, but if a small particle of foreign matter is introduced, heat is evolved and the temperature rises to 32·4° C.—the normal fusion-point. The phenomenon of supercooling can therefore be used in the accurate determination of the melting-point of a substance. DESPRETZ noticed the same effect in capillary tubes containing water and it has been suggested that this is possibly the reason why the sap often remains liquid in the capillary vessels of plants during a spell of cold weather.

The Change in Volume on Solidification.—In general, a contraction occurs when a substance passes from the liquid to the solid state. This may be shown by melting tin in a crucible and then allowing it to cool. When the tin is cold a distinct cavity will be seen where the surface was flat when the tin was liquid. Paraffin wax behaves in the same way. Water, bismuth, and antimony exhibit the reverse effect. The increase in volume when water freezes may be shown with the apparatus indicated in Fig. 11·2. A large test-tube is half-filled with water and turpentine poured in on the water so that the tube is filled completely. The tube is closed with a cork provided with a long piece of glass tubing. After it has been ascertained that there are no air bubbles present, the tube is surrounded by a freezing mixture [cf. p. 241]. At first the level of the oil may fall slightly due to a decrease in temperature, but soon the oil will rise rapidly. On removing the tube from the freezing

mixture ice will be noticed in it and it is the expansion due to its formation that causes the oil to rise.

The expansion accompanying the production of ice brings in its train many beneficial results, but unfortunately also some that are destructive. Amongst the beneficial results it may be mentioned that the fertility of soil is increased because of the disintegration of its parts during frosty weather. The same action is responsible for the weathering of rocks. Moreover, life, as we know it on this planet, is only possible because of this expansion. Had it been otherwise the ice formed in ponds and streams during the winter would sink, and the heat of summer would not be sufficient to melt it. Year after year conditions would pass from bad to worse until all life depending upon an adequate supply of water would become extinct.

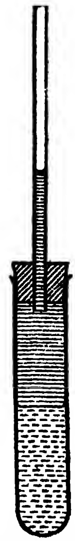


FIG. 11·2.

On the other hand, this same expansion causes water-pipes to burst, and is often sufficient to raise a pavement. To prevent lead pipes from bursting it has been proposed to make them square in section so that when ice formed only the shape of the section would alter—it would tend to become circular. The proposal has not been adopted on account of the difficulties of bending such tubing and of making a T-joint in it.

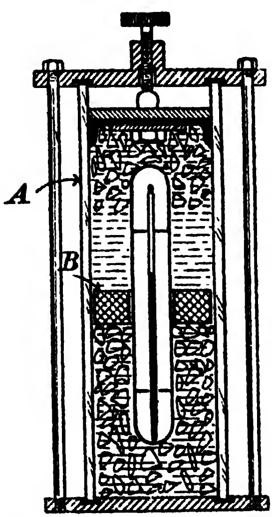


FIG. 11·3.—Apparatus for Investigating Effect of Pressure on Melting-points.

Influence of Pressure on the Melting-point.—We have already indicated that the melting-point of a substance depends upon the pressure to which the substance is subjected. JAMES THOMSON, in 1849, first showed that the melting-point of a substance which expands on solidifying [ice] would be lowered when the pressure was increased. Simple reasoning shows that this may be expected, for if a substance expands on solidification increased pressure will be unfavourable to such a change. Thomson calculated that the melting-point of ice would be lowered by 0·0075° C. per one atmosphere increase, i.e. *in vacuo*, ice melts at + 0·0075° C. Thomson's brother, the late LORD KELVIN, devised the following experiment to test this conclusion. His

apparatus, Fig. 11·3, consisted of a strong glass cylinder, A, containing ice and water. A thermometer was placed inside and a massive piece of lead, B, in the form of a ring kept the central portion of the apparatus free from ice so that the indications of the thermometer could be observed. The thermometer contained a mixture of ether and sulphuric acid as the thermometric substance. A glass case protected the thermometer from the effects of increased pressure which would have tended to make the thermometer reading too high, due to a diminution in the volume of the bulb when the pressure was increased. The bulb of the thermometer was large so that the instrument was sensitive but slow in action. [At a later date Callendar used a platinum resistance thermometer which, in addition to being sensitive, was unaffected by changes in pressure and more rapid.] The outer glass vessel was closed by a metal lid provided with a screw plunger. By rotating the screw the pressure inside the apparatus could be increased. The pressure was measured by noting the compression of air enclosed in a vertical tube not shown in the diagram.

FIG. 11·4.—Apparatus for Investigating Effect of Pressure on Melting-points.

To study the effect of pressure on the melting-point of wax, a substance contracting on solidification, BUNSEN devised the apparatus shown in Fig. 11·4. The shorter arm AB contained wax whilst air filled the portion DE. The bulb C and the rest of the apparatus were filled with mercury. The temperature of C was increased by heating the water bath around it. This caused the mercury to expand and exert a considerable pressure on the wax. The actual expansion was controlled by inserting the apparatus to different depths in the bath. The pressure was deduced from observations on the volume of the air above D. Bunsen found that paraffin wax, melting at 46·3° C. under a pressure of one atmosphere, melted at 49·9° C. when subjected to a pressure of 100 atmospheres.

Regelation.—If a thin loop of copper wire from which a heavy load is suspended is passed round the middle of a large block of ice at 0° C., the pressure on the ice under the wire is considerable so that the ice melts and the wire passes into the block. The water thus formed passes round the wire and freezes again since the excess pressure upon it has been removed. This process continues until

after several hours the wire will have cut its way through the ice, but the block will remain intact. An important process persisting throughout the whole operation just described is the constant flow of heat through the copper wire. The water immediately behind the wire is solidifying at its normal freezing-point, whereas the ice underneath is melting at a temperature lower than 0° C., owing to the pressure to which it is subjected being much greater than atmospheric. Now the solidification of the water above the wire is accompanied by an evolution of heat while the melting of the ice below necessitates an absorption of heat. There will therefore be a flow of heat downwards through the wire and this maintains both actions in process at the same time. From this it is clear that the greater the thermal conductivity of the wire, the more quickly will the latter cut its way through the ice. As an extreme illustration of this the wire may be replaced by catgut when the cutting process is greatly retarded and the water does not freeze above the catgut.

The melting of ice under increased pressure is shown by the following experiment. Two blocks of ice adhere when pressed together and the pressure removed afterwards. This adhesion even occurs if the experiment is repeated with the ice blocks immersed in warm water.

The slippery nature of ice is also due to the fact that ice melts more easily under increased pressure—very cold ice is not slippery, and, for the same reason, very cold snow cannot be formed into a snowball.

The motion of a glacier is attributed, at least in part, to this same phenomenon. The snow which accumulates to immense depths on high mountains exerts an enormous pressure on the underlying masses, which melt, and then freeze into solid ice when the pressure is removed. The increased pressure of the snow and ice at the source causes the lower strata of the glacier, which was thus gradually formed, to melt and solidify alternately. At each melting the glacier moves forward. This process is referred to as ***regelation***.

The Freezing of Solutions.—Experiment shows that the freezing-point of a solution is lower than that of the pure solvent. As an example let us consider the effect of lowering the temperature of an aqueous solution of sodium chloride. It is found that the temperature at which ice begins to form and separate out from such a solution decreases as the proportion of salt is increased. This continues until a temperature of − 23° C. has been obtained. Any further decrease in temperature and the whole solidifies *en bloc*. The solution which solidifies at this temperature contains 23·6 per cent. of salt. If solutions having a higher concentration of salt than this are cooled it is found that the freezing-point decreases as the amount of salt ***decreases***, i.e. as the proportion of water ***increases***.

Moreover, it is salt and not ice which now separates when the freezing-point is reached. If the solution contains exactly 23·6 per cent. of salt then neither ice nor salt is deposited as the solution cools but the whole solidifies at a temperature of − 23° C. This is termed the ***eutectic temperature,*** while the particular mixture which freezes at this temperature is known as the ***eutectic.***

The Fusion of Alloys.—As a particular example let us consider the thermal equilibrium of alloys of thallium and gold. The freezing-point of pure metallic thallium is 300° C. and is represented by A, Fig. 11·5. As the gold content increases the freezing-point moves along AB until the eutectic point B is reached. Similarly, pure gold melts at 1100° C.—the point C on the diagram. When thallium is added to gold the temperature of solidification falls until the eutectic composition is again reached.

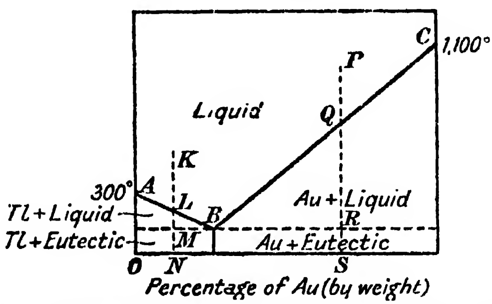


FIG. 11·5.—Thermal Diagram for Alloys of Thallium and Gold.

Consider what happens when an alloy containing ON per cent. gold is cooled from the temperature indicated by the point K. The whole remains liquid until L is reached, when thallium begins to separate. The amount of thallium thus separating increases until the temperature corresponds to M. The mother liquid then has the eutectic composition so that after passing M all is solid—an intimate conglomerate of thallium and the thallium-gold eutectic. Similarly, if the alloy contains OS per cent. gold and its initial temperature is P everything remains liquid until Q is reached, when gold begins to crystallize out. The amount of gold increases until R is reached after which the whole is solid. This consists of the thallium gold eutectic throughout which pure gold is distributed. This particular example has been chosen since AB and BC are nearly straight lines.

The alloys of tin and magnesium are much more interesting from this point of view. The thermal diagram is given in Fig. 11·6. It is at once apparent that there is a definite maximum on the curve corresponding to an alloy containing 30 per cent. magnesium by weight. This corresponds to the intermetallic compound Mg_2Sn.

The curve exhibits two eutectic points, one for the system tin and Mg_2Sn, and the other for the system magnesium and Mg_2Sn. The effect of cooling any particular alloy in this series is shown by the lettering in the diagram.

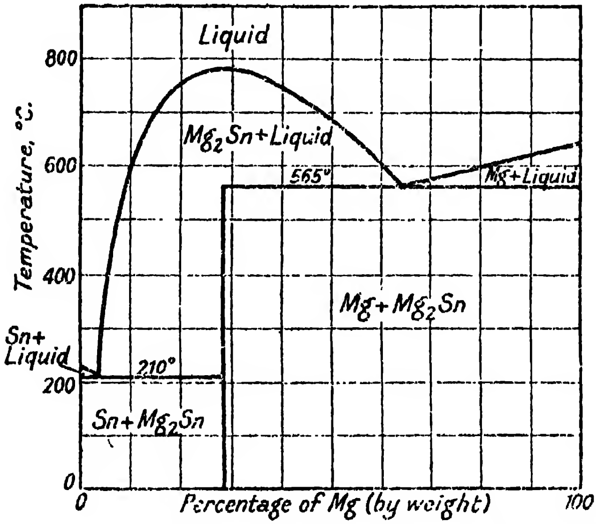


FIG. 11·6.—Thermal Diagram for Alloys of Magnesium and Tin.

Low Melting Point Alloys.—(i) ***Wood's metal.***—This is an alloy of tin, lead, cadmium and bismuth, in the proportions 1 : 2 : 1 : 4. It melts at 60·5° C.

(ii) ***Rose's metal.*** Tin, 1; lead, 1; bismuth, 2. Melting-point 94·5° C.

These readily fusible alloys find many applications in daily life, e.g. in automatic sprinklers for buildings, so that when a fire occurs a plug made from one of these alloys and inserted in a water pipe melts and the water rushes from the mains. Also, a flow of gas along a pipe may be stopped if pieces of such alloys have been placed in the pipe, for they melt when a fire breaks out. Fusible plugs also permit fireproof doors to close automatically in the event of a fire. These alloys are also used as fuses in electrical circuits.

EXAMPLE XI

1.—Describe and explain an experiment by means of which the effect of increased pressure on the melting point of ice may be investigated. How may the motion of a glacier be explained?

CHAPTER XII

EVAPORATION AND EBULLITION—THE PROPERTIES OF VAPOURS

The Three States of Matter.—All substances are composed of molecules, or groups of atoms, and in solids these particles are held together by large forces—it is said that solid bodies possess the property of cohesion. There is no reason to believe, however, that these particles are at absolute rest; in fact, if a piece of lead is placed in contact with gold and left for a period of several years, gold atoms are found embedded in the lead, showing that a shifting of the atoms has taken place. When the above experiment is performed at higher temperatures the migration of the atoms is facilitated. From such facts one must conclude that the molecules of a substance have a velocity which increases with rise in temperature. If the temperature of a solid is raised continuously, a stage is reached when the binding forces between the molecules cease to be sufficient to restrain their motion to the same extent as in a solid. The solid has changed into a liquid. A further supply of heat to the liquid again diminishes the forces of cohesion until the liquid becomes a vapour: in this last state of matter the molecules are more free to move than in any other.

It must not be supposed that all the molecules in a body at a fixed temperature possess the same velocity; in fact, for gases, Maxwell was able to calculate what fraction of the molecules in a gas were moving with a velocity different from that of the majority of the molecules. In liquids, for example, there will be some particles moving with a velocity greater than that of the major portion of the molecules, and these will, of course, move about more easily. Should they happen to be at the surface of the liquid, where, as mentioned in the section on surface tension, the resultant force is directed inwards, then the velocity being large, the energy of the molecules may be sufficiently big for them to overcome this force and to wander outside the realm of attraction of the liquid. Fig. 12·1 is a diagrammatic representation of the state of affairs near the surface of a liquid in contact with air. The trajectories of the liquid molecules are indicated. Some of the molecules do not possess sufficient energy to escape from the liquid completely; they pass into a

region just beyond the liquid and then return to it. They have merely made a transient excursion into the so-called 'region of molecular attraction.' Other molecules, possessing more kinetic energy, leave the liquid and do not return: the liquid is evaporating. Now an increase in the temperature of a liquid is accompanied by an increase in the mean kinetic energy of its molecules, so that more escape from the liquid—the rate of evaporation has been increased. [A current of air also facilitates evaporation since some of the molecules not passing normally beyond the confines of molecular attraction are removed.]

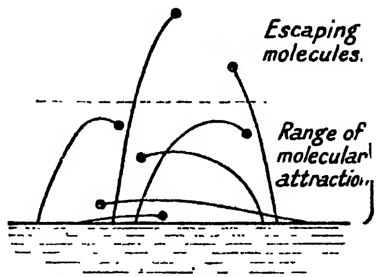


FIG. 12·1.—Diagrammatic Representation of Molecules escaping from the Surface of a Liquid.

The molecules escaping from the liquid in the form of a vapour exert a pressure in just the same way as do the molecules of a gas, since they possess momentum.

The Saturation Vapour Pressure of a Liquid.—Let a small quantity of the liquid be introduced into the Torricellian vacuum of a barometer tube—Fig. 12·2. The introduction of the liquid is facilitated by means of the small pipette; in this operation it is advisable not to use the lungs to apply the necessary pressure, it being better to attach a piece of rubber tubing to the end of the pipette. This tubing is closed with a short piece of glass rod, and then squeezed so that the pressure inside the pipette increases and causes the liquid to be exuded and to rise above the mercury. At first, if a sufficiently small quantity of liquid has been used, the mercury column is depressed but no liquid is visible in the tube. The molecules of the liquid have escaped into the previously exhausted space, which is now said to contain an ***unsaturated vapour***. By continuing the process, the mercury is further depressed, until suddenly there appears a small quantity of liquid on the surface of the mercury. When this happens just as many molecules leave the surface as return to it—the equilibrium, however, is a dynamic one, for there is no reason to suppose that the molecules of the vapour or the liquid have ceased to move when this condition is attained. The space above the mercury is now saturated with vapour, and the depression of the mercury is equal to the ***saturation vapour pressure*** of the liquid at the temperature of the experiment, expressed in terms of centimetres of mercury.

If the barometer tube is supported with its open end below mercury contained in a deep vessel so that the volume of the space above the mercury in the tube may be varied, it will be found that the height of the mercury in the tube remains constant, so long as there is liquid remaining in contact with the vapour. The volume of the

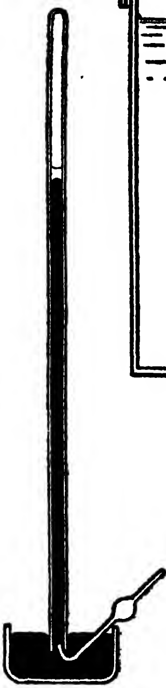


FIG. 12·2.

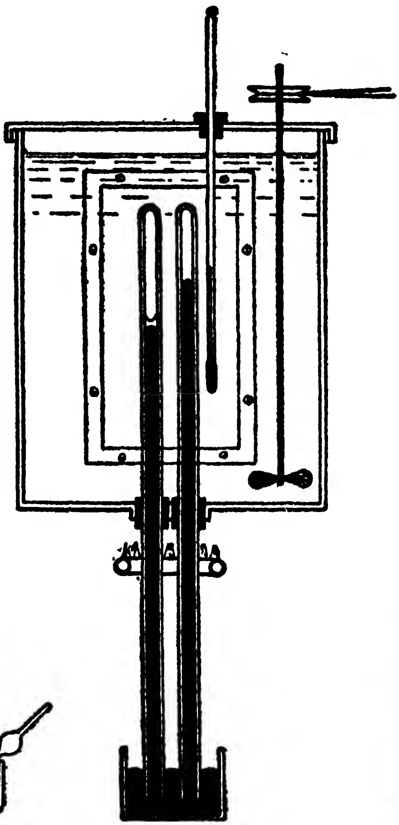


FIG. 12·3.

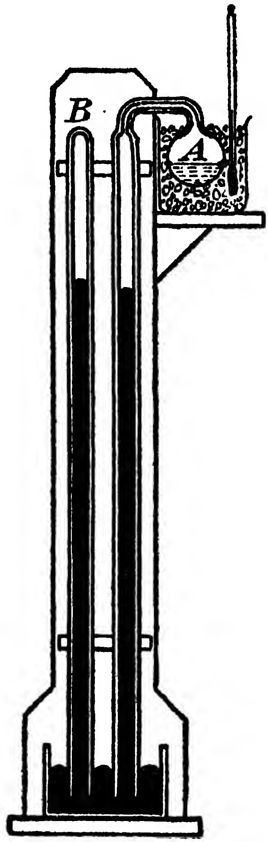


FIG. 12·4.

FIG. 12·3. Regnault's Apparatus for S.V.P. of a Liquid when the S.V.P. does not exceed 30 cm. of Mercury.

FIG. 12·4.—Regnault's Apparatus for S.V.P. of Water below 0° C.

vapour may increase, for example, but its pressure is unaltered—during the process, however, some of the liquid will evaporate; if the volume is decreased, condensation ensues—by which process the pressure is maintained constant.

Saturated vapours, i.e. vapours in contact with their own liquids,

are characterized by the fact that, at constant temperature, if the volume changes the pressure remains unaltered.

DALTON performed experiments similar to the above and was able to enunciate two laws :—

(1) *The pressure exerted by a saturated vapour depends only upon the temperature and the particular liquid used.*

(2) *The pressure exerted by a mixture of vapours (or gases) is equal to the sum of the pressures which each would separately exert if it alone occupied the space filled by the mixture.* [Dalton's law of partial pressures.]

Regnault's Apparatus for the Measurement of the S.V.P. of Water and Other Liquids at Low Temperatures.—This is a modification of an earlier form of apparatus designed by DALTON. The upper portions of two barometer tubes, Fig. 12·3, were placed in a bath which could be heated. The bath was well stirred and was made large [45–50 litres] to minimize fluctuations in temperature. Water was introduced into one tube, the quantity being controlled so that the space above the mercury was saturated with vapour. The difference in level between the two mercury surfaces, observed by a cathetometer, gave the saturation vapour pressure at the temperature of the bath in terms of cm. of mercury at the same temperature. To get comparable values at different temperatures the results were corrected to 0° C. The front of the bath was provided with a plate-glass window. The observations were repeated at other temperatures. Regnault found that reliable results could not be obtained at temperatures above 50° C. since the bath became too long for its temperature to be kept constant. Moreover, when the mercury was depressed nearly the whole length of the bath there was a tendency for the mercury surface to be cooler than the bath, so that it became difficult to know the true temperature corresponding to the pressure measured.

Regnault's Apparatus for Water below 0° C.—The apparatus consisted of two barometer tubes, Fig. 12·4, to one of which a bulb, A, had been sealed. This contained water and was surrounded by a freezing mixture of snow and calcium chloride. The depression of the mercury in the experimental tube below that in B is due to the pressure of the vapour [1] in A, and by measuring this depression the vapour pressure was determined.

[1] In the study of vapours there are many pitfalls, a very common one being associated with the following experiment :—Suppose a glass cylinder is hermetically sealed and contains *only* a liquid (say ether) and its vapour. The ether is immersed in ice ; this condition is sufficient to determine the pressure inside the apparatus provided that the temperature is everywhere not less than 0° C. If the upper region of the tube is gently heated some ether will condense, but the pressure inside the apparatus is still that of the vapour pressure of ether at 0° C. If two such pieces of apparatus at

Regnault's Apparatus for Water at Higher Temperatures. —We have already noticed the objections to Regnault's first apparatus when an attempt was made to use it at higher temperatures. To overcome these, REGNAULT built the apparatus depicted in Fig. 12·5. The method is based on the fact that when a liquid boils its saturation vapour pressure is equal to the pressure of the 'atmosphere' in which the boiling occurs—cf. p. 252. The water was heated in a copper vessel, B, connected to a bulb, A, containing air. The tube connecting A and B sloped upwards and was surrounded by a condenser, C. The air could be removed in part from

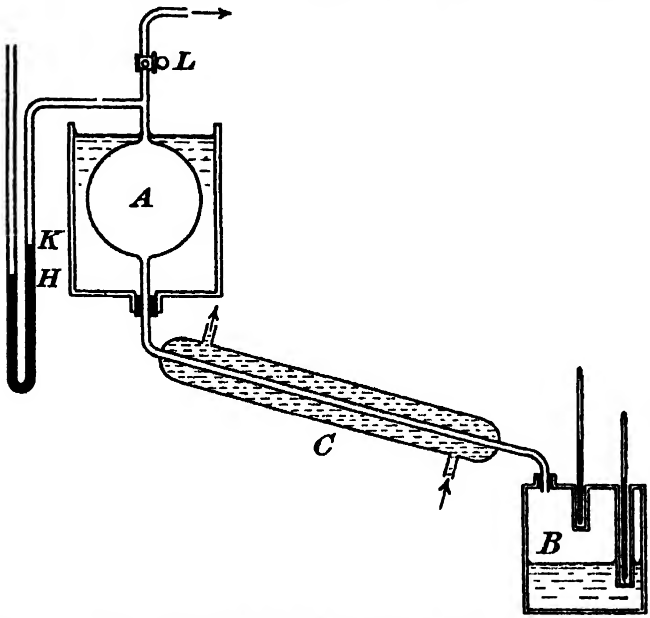


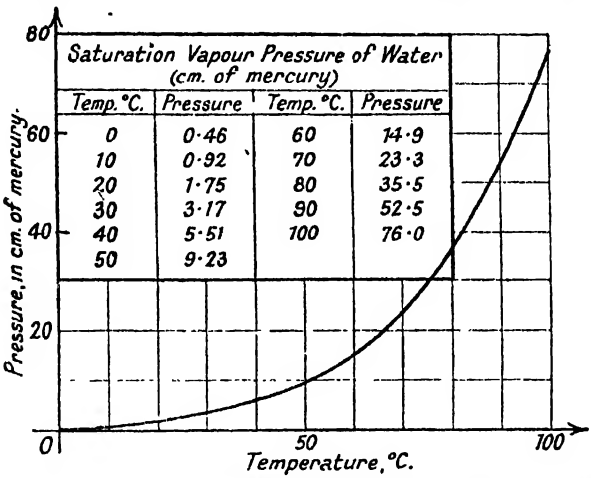
FIG. 12·5.—Regnault's Apparatus for S.V.P. at Higher Temperatures.

the apparatus by a suitable pump. The pressure in the bulb was atmospheric less that due to a column of mercury HK in the manometer. Four thermometers gave the temperature in B—only two are shown here. They were inserted in small cavities closed at their lower ends and containing mercury to give good thermal contact. Two thermometers were in the vapour and two in the liquid. Their readings were the same, indicating an

0° C. and 20° C., respectively, are connected together by means of a glass tap, then, when the tap is opened, the pressure is at once everywhere equal to that of ether at 0° C. : but ether at 20° C. exerts a greater pressure than this ; so the ether at 20° C. evaporates and condenses in the other tube. This experiment shows that the vapour pressure above a liquid is always equal to the vapour pressure of the liquid at the lowest temperature existing.

absence of delayed boiling. The pressure having been adjusted to some desired value, the tap L was closed and in a short time the thermometers indicated a steady temperature; the vapour pressure was then deduced since, because the liquid is boiling, it is the pressure of the atmosphere in the apparatus. By placing the thermometers in cavities as shown and not directly in contact with the liquid or its vapour, Regnault avoided errors due to the effects of a varying pressure on the bulbs of the thermometers.

Although we have always referred to water as the liquid under examination it is at once apparent that the saturation vapour pressures of all other liquids which do not react with mercury or glass can be determined by one or other of these methods. This



Saturation Vapour Pressure of Water (cm. of mercury)			
Temp. °C.	Pressure	Temp. °C.	Pressure
0	0·46	60	14·9
10	0·92	70	23·3
20	1·75	80	35·5
30	3·17	90	52·5
40	5·51	100	76·0
50	9·23		

FIG. 12·6.—Saturation Vapour Pressure of Water at different Temperatures.

apparatus may also be employed to determine the S.V.P. of a liquid at temperatures above its normal boiling-point, when the only change in procedure is to increase the pressure of the air in the apparatus. If the effect of increasing the pressure considerably above atmospheric is being investigated, a closed manometer must be used to measure the pressure inside the apparatus. The pressure is deduced from observations on the volume of air—or better nitrogen—contained in the closed limb of the instrument. Moreover, the various joints must be suitably strengthened.

The manner in which the saturation vapour pressure of water varies with temperature is shown in Fig. 12·6.

Ramsay and Young's Apparatus.—RAMSAY and YOUNG (1885) devised a convenient and accurate method of determining the vapour pressure of a liquid at temperatures such that the saturation vapour pressure does not exceed 50 cm. of mercury. The arrangement is

shown in Fig. 12·7. A glass tube A, with a side tube B, carries a thermometer T, and a dropping funnel C containing the liquid. The side tube B leads to a small bottle D, and this leads to a suitable manometer. The pressure of the air in A is controlled by a pump connected to E. The bulb of the thermometer T is surrounded with cotton or asbestos wool on to which liquid from the funnel is caused to drop, this being facilitated by the bend at the lower end of the dropping funnel.

The tube A is placed in a suitable oil bath so that its temperature may be raised. Liquid is then allowed to flow on to the cotton-wool attached to the bulb of the thermometer until the wool is thoroughly wetted. Rapid evaporation ensues and the vapour displaces the air in the lower part of the tube A; in this region the liquid on the wool

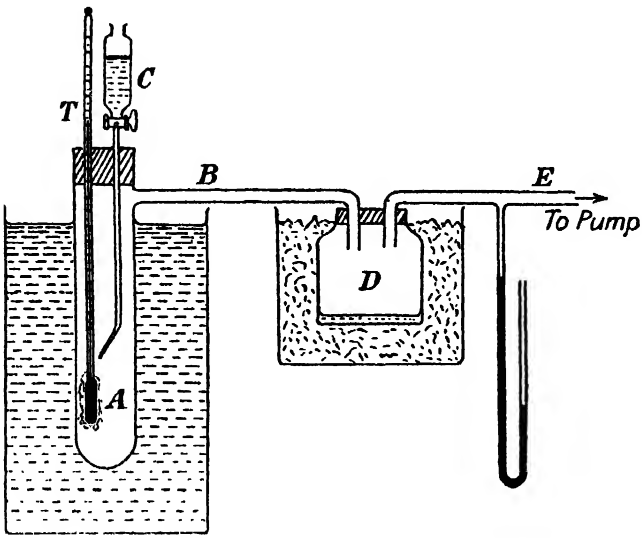


FIG. 12·7.—Ramsay and Young's Method of Determining Saturation Vapour Pressures.

is surrounded by an atmosphere of its own vapour in which there is very little air. Under these circumstances the liquid soon reaches a state in which it is in equilibrium with its own vapour, i.e. it has reached its boiling-point for the particular pressure to which it is subjected: free-boiling is impossible, and the vapour cannot be superheated as long as any liquid remains on the cotton-wool, but as the vapour gradually diffuses away towards D, further evaporation follows. The pressure inside the apparatus is equal to the saturation vapour pressure of the liquid at the temperature indicated by T. It must be emphasized that it is only round the bulb of the thermometer that there is a saturated vapour—towards D, a vessel surrounded by ice, there is a mixture of air and vapour but beyond D there is only air, if D is efficient in condensing the vapour which enters that vessel by diffusion. The pressure in the apparatus, however, is constant and equal to the saturation vapour pressure of the liquid at the temperature recorded by T. When the thermometer shows a steady temperature the reading on the

manometer is recorded. This pressure difference, subtracted from atmospheric pressure, gives the vapour pressure of the liquid at the temperature indicated by T. Consistent results are obtained when the temperature of the oil bath is about 20° C. above that of the steady temperature indicated by T.

The above apparatus was used by Ramsay and Young to determine the saturation vapour pressures of camphor and acetic acid. They found that their results were most concordant for pressures not exceeding 50 cm. of mercury, i.e. for temperatures below and not too close to the normal boiling-point of the liquid investigated.

Determination of the S.V.P. of Bromine.—The saturation vapour pressure of a liquid [say bromine] which attacks mercury may be determined as follows :—A long glass bulb, B, Fig. 12·8, is blown and made into a form of hollow spoon by heating the glass on one side and applying suction while the bulb is hot. A long light glass pointer is attached to B. This portion of the apparatus is surrounded by a wide glass tube, the pressure of the air in it being controlled by a pump connected above the tap, C. The pressure is recorded by the manometer DE in which the space above E is exhausted. The position of the end of the pointer on the scale S is noted when A and the tube around B are exhausted. Bromine is then introduced into A which is surrounded by a water bath and, providing that the temperature of B and the tube connecting it to A is greater than that of A, the pressure in A, and therefore in B, is equal to the saturation pressure of bromine at the temperature of A. The excess temperature just mentioned is obtained by a heating coil placed as shown. The pressure in the wide tube is then adjusted so that the end of the pointer which has become deflected in this process is brought back to its zero position. The pressure of the air in this tube is equal to the pressure in A.

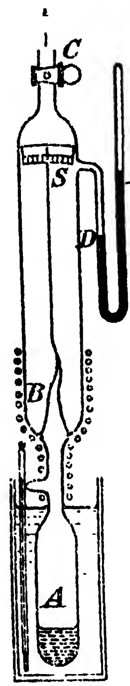


FIG. 12·8.—S.V.P. of Bromine.

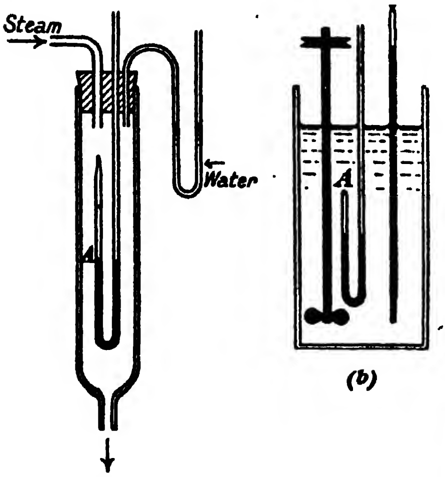
Conditions under which Boiling Occurs.—Before a liquid boils small bubbles of vapour are seen in those portions of the liquid nearer to the supply of heat: if these are watched carefully it will be noticed that they disappear before reaching the surface of the liquid. The latent heat given out in this process helps to warm the upper regions so that eventually the temperature becomes uniform throughout the liquid, and the liquid boils.

It should be noted that liquids evaporate when exposed to the atmosphere at a rate which is greatly accelerated by increase in temperature; on the other hand, boiling only occurs when the S.V.P. of the liquid equals atmospheric pressure.

Steady Boiling.—Some time after a liquid has commenced to boil it may sometimes be noticed that 'explosive boiling' or

'boiling by bumping' occurs. This is attributed to the fact that the nuclei necessary for steady boiling have disappeared or become inactive. Steady boiling may be re-established by introducing a few fragments of broken glass or porous porcelain. The gas enclosed in the material will maintain steady boiling for a considerable time, especially if the supply of heat is not too vigorous.

The S.V.P. of a Liquid at its Normal Boiling-point.—A U-tube, closed at one end, is made and completely filled with mercury by the method of alternate heating and cooling. A little mercury is removed and replaced by recently boiled distilled water [air free]; by inverting the tube, having closed the open end with the first finger, the water is introduced into the closed limb of the tube. All the mercury, except for a length of a few cm., is removed from the



(a)

FIG. 12·9.

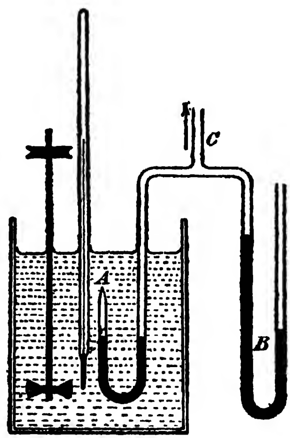


FIG. 12·10.—Variation of B.P. of a Liquid with Pressure.

open limb—this may be done by means of a capillary tube, drawn out from wider tubing so that suction may easily be applied. The tube is then placed in steam, Fig. 12·9 (*a*), when it will be found that the mercury stands at the same level in each limb of the U-tube. This experiment shows that the S.V.P. of water when it boils under atmospheric pressure is equal to the pressure of the atmosphere.

If it is necessary to find the boiling-point of a liquid, especially if it is available in small quantity only, another tube similar to that just described is made, only the liquid under examination is introduced instead of water. The whole is placed vertically in a bath containing liquid which boils at a higher temperature than that being investigated, and the temperature raised until the mercury

is at the same height in each limb, the bath being well stirred. The temperature of the bath is noted; the experiment is repeated with the temperature of the bath falling; the mean of these is the boiling-point required—cf. Fig. 12·7 (*b*). In these experiments due caution must be exercised to see that a little liquid still remains in the limb A.

The S.V.P. of a Liquid (Small Quantity Available) and its Variation with Temperature.—The apparatus is shown in Fig. 12·10. It consists of a U-tube whose closed limb A contains the liquid above mercury. The other end of this tube is connected to a second U-tube B containing mercury. The pressure inside the apparatus may be varied by connecting the tube C to a pump. The boiling-point of the liquid at different pressures is then investigated by inserting A in a water or oil bath and proceeding as in the previous experiment.

Vapour Pressure of Solutions.—Experiment reveals the fact that the saturation vapour pressure of a solution is less than that of the pure solvent. It therefore follows that, when such a solution is at the temperature at which the solvent would boil under the prevailing conditions, the vapour pressure of the solution is less than atmospheric pressure, so that the solution does not boil: it only boils when the temperature is raised above this value. To determine the boiling-point of a solution the thermometer must be placed in the liquid. The reason for this is that if we are dealing with an aqueous solution, for example, the steam from the liquid would condense on the thermometer bulb which would indicate a temperature corresponding to the steam temperature under existing circumstances. Any further heat supplied to this water simply causes it to evaporate without increasing its temperature, and since this supply comes from steam at a slightly higher temperature, water will condense on the bulb as fast as it evaporates away. Steady boiling is maintained by one of the methods already described.

Variation of Boiling-point with Pressure.—We have already learnt that when a liquid boils its saturation vapour pressure is equal to that external pressure acting on its surface. It therefore follows that if the variation of the saturation vapour pressure with temperature be known the boiling-point of the liquid at different pressures is also known. Regnault's apparatus may be used to investigate this effect. By abstracting air from the apparatus the boiling-point at pressures less than one atmosphere may be found, while by increasing the pressure of the air the boiling-point at higher pressures may be found. It is interesting to note that Regnault himself used this apparatus up to a pressure of 28 atmospheres and proposed to construct a stronger apparatus to withstand greater pressures.

The effect of reduced pressure on the boiling-point of water may be

shown in a very striking way as follows :—A round-bottomed flask [any other shaped flask invariably cracks owing to the increased strain at any corner when the pressure outside differs from that inside] is half-filled with water which is boiled to expel most of the air. While steam is still issuing, the mouth of the flask is closed with a rubber bung, the flame removed *at once,* and the flask inverted under a stream of cold water. The water continues to boil vigorously, even though the temperature is reduced as low as 40° C.

The fact that liquids boil at lower temperatures when the external pressure is reduced is often used in the manufacture of certain classes of substance, especially those decomposing at a higher temperature. For example, milk is boiled under such conditions in the production of milk powder ; sugar is also refined by a similar process.

The Hypsometer.—The atmospheric pressure at any station is equal to the weight of a column of air, 1 cm.^{2} in section, stretching from that station to the upper limit of the atmosphere. It therefore follows that at high altitudes the pressure is less than at sea-level. If the difference in pressure between two stations is known, the difference in their altitude may be derived at once if the density of air is also known. To carry a barometer from place to place would be cumbersome, so that it is preferable to use indirect means of ascertaining the pressure. This can be done with the aid of a ***hypsometer,*** an instrument employed in measuring the boiling-point of water at pressures other than that of the standard atmosphere [when it is defined as 100° C.]. This instrument resembles that used in discovering the error of a mercury thermometer at the upper fixed point. The thermometer, however, differs in one respect from the usual mercury-in-glass thermometer. The mercury column is broken near the upper end by a bubble of air. In using the instrument this short piece of mercury is shaken down ; after use the thermometer is removed from the hypsometer when the thread still remains in position, thus indicating the maximum temperature to which it has been subjected.

Dalton's Law for Mixed Vapours.—For a mixture of two or more gases or vapours which do not react chemically with one another Dalton discovered that the total pressure was equal to the sum of the pressures that each component would exert if it were present alone and occupied the same volume as does the mixture. This is known as ***Dalton's law of partial pressures.*** It can, of course, only be an approximation since it is impossible to establish an infinitely great pressure by mixing a large number of vapours. But as a first approximation it is true both for saturated and unsaturated vapours.

Example.—Moist oxygen is confined over water at 20° C. The total pressure is 758·2 mm. of mercury; if the saturation vapour pressure of water at 20° C. is 17·4 mm. of mercury, calculate the pressure of the oxygen alone. From Dalton's law of partial pressures it follows at once that the partial pressure of the oxygen is (758·2 − 17·4) = 740·8 mm. of mercury.

Experiment.—Introduce a small quantity of air into a barometer tube containing mercury. Let the column be depressed x cm. so that the pressure of the air inside the tube is x cm. of mercury; let the length of tube occupied by the air be l_1 cm. Suppose that when a liquid is introduced into the space above the mercury the total depression is y cm., l_2 being the length of the tube occupied by the mixture of air and vapour. The air in the tube is now exerting a partial pressure p given by

$$l_1ax = l_2ap$$

where a is the cross-section of the tube. The partial pressure due to the liquid is therefore $(y - p) = \left(y - \frac{l_1}{l_2}x\right)$ cm. of mercury.

If a little liquid remains as liquid in the tube the above is the saturation vapour pressure of the liquid at the temperature of the experiment.

Experiment.—If the saturation vapour pressure of water at 20° C. is known, its value at another temperature—say 50° C.—may be determined as follows. A water index about 2 cm. long is used to enclose a volume of air in a capillary tube of uniform diameter. The tube is placed in a vertical position in a well-stirred bath of water at 20° C., and l_{20}, the length of the tube occupied by the air which is saturated with water vapour at this temperature is determined. The temperature of the bath is raised to 50° C. and l_{50} determined. Let p_{20} and p_{50} be the saturated vapour pressures of water at 20° C. and 50° C. respectively. If P is the atmospheric pressure and therefore the total pressure inside the tube, the partial pressures of the air at these two temperatures are $(P - p_{20})$ and $(P - p_{50})$ respectively. Applying the laws of Boyle and Gay-Lussac to the air, we have

$$\frac{(P - p_{20})l_{20}}{T_{20}} = \frac{(P - p_{50})l_{50}}{T_{50}},$$

so that p_{50} may be determined if p_{20} is known.

Experiment.—Introduce a little water above the mercury in the closed limb of a Boyle's Law tube and make a series of observations of corresponding values of the pressure and volume of the mixture of air and saturated vapour. Plot the values of the pressures as ordinates against the reciprocals of the corresponding volumes as abscissæ. Draw the best straight line through the points thus obtained. The intercept made by this line on the y-axis is the saturation vapour pressure of water at the temperature of the experiment. The reason for this is that if P is the pressure of the air and vapour, p the S.V.P. of the water, then, considering the air alone, $(P - p)V$ = constant (say a), or $P - p = a/V$. Hence, if $y = P$ and $x = \frac{1}{V}$, this equation becomes $y = ax + p$, which represents a straight line whose intercept on the y-axis is p, the required pressure.

The Triple Point.—In Fig. 12·11, the curve OP represents the relation between the vapour pressure of water (liquid) and its temperature: it is termed the ***steam line.*** The vapour pressure curve for ice is represented by OQ, a curve known as the ***hoar-frost*** or ***sublimation curve.*** The effect of pressure on the melting-point of ice is represented by the line OR, the so-called ***ice-line.*** The point O, where the three lines intersect is known as the ***triple point,*** its co-ordinates being pressure 4·57 mm. of mercury, temperature 0·0075° C. This point is such that at the pressure and temperature represented by it, the three phases, solid, liquid and vapour may co-exist in equilibrium. Any small departure from these conditions is accompanied by the disappearance of one of the phases and the equilibrium is represented by a point on one of the lines, OP, OQ, or OR.

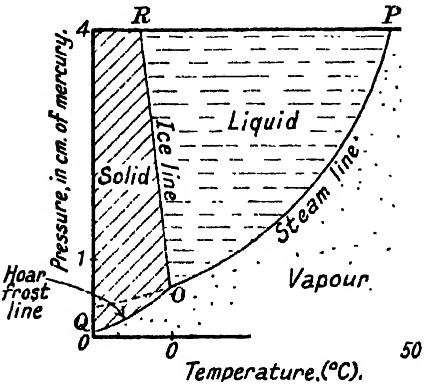


FIG. 12·11.—The Triple-Point for Water.

Definitions: The components of a system (of substances) are those substances taking part in a reaction but not decomposed in the process. The components may be elements or compounds—in the instance discussed above there is one component, viz., water.

The phases of a system are the different physical states in which the components may exist. Thus ice, liquid water, and vapour are the three phases in which water may exist.

Returning to Fig. 12·11, it may now be said that the lines OP, OQ, and OR represent the equilibrium conditions of three two-phase systems: liquid-vapour, vapour-solid, and solid-liquid.

Vapour Density.—The vapour density of a substance is defined as the density, i.e. mass per unit volume, which the vapour would possess if it could exist as an ideal gas at 0° C. and under a pressure of 76 cm. of mercury. Although such conditions can never be realized, it is usual to make the calculation as if such conditions were possible. In other words, observations on the pressure and volume of a given mass of vapour at the steam temperature are made, and the vapour is then treated as an ideal gas in order to calculate the volume at S.T.P. Let V be the volume of the vapour

[unsaturated], at pressure p and temperature T° K. Then at 0° C. [or 273° K.] and 76 cm. of mercury pressure, its volume V_0, in in so far as the vapour may be considered an ideal gas, is given by

$$\frac{76\,V_0}{273} = \frac{pV}{T}, \text{ or } V_0 = \frac{p}{76} \cdot \frac{273}{T} \cdot V.$$

If m is the mass of liquid used, its vapour density ρ, as defined above, is

$$\rho = \frac{m}{V_0} = m \cdot \frac{76}{p} \cdot \frac{T}{273} \cdot \frac{1}{V} \text{ gm. cm.}^{-3}$$

One of the most accurate means of determining the vapour density of a substance is due to HOFMANN. The apparatus is shown in Fig. 12·12. A glass tube A, having a diameter of 2 cm. and length 90 cm., is cleaned, dried, filled with pure clean mercury, and finally inverted in a trough B containing mercury. The tube A is surrounded by a wider jacket C, through which steam or other vapour may be passed. The two tubes dip into B, the tube A being held in position by means of a circular piece of wire gauze D.

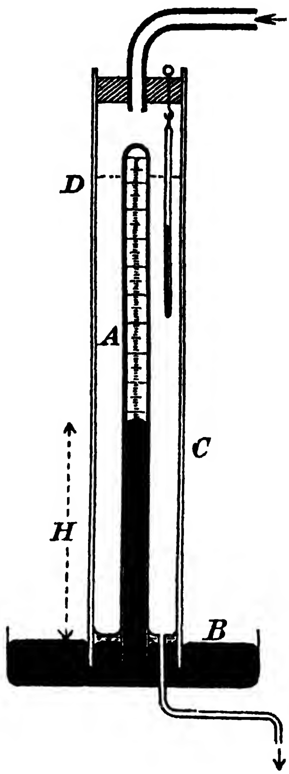


FIG. 12·12.—Hofmann's Vapour Density Apparatus.

[N.B. The thermometer shown suspended above is not required if steam is the only vapour used to vaporize the liquid under investigation.]

A small known mass of the liquid whose vapour density is required is placed in a miniature bottle and inserted under the barometer column; the bottle rises and the liquid evaporates, causing the mercury column to be forced downwards—of course, all the liquid must evaporate. The volume of the vapour is deduced from the graduations on the tube, parallax errors being avoided by placing a mirror behind the tube, whilst the pressure is obtained as follows. Let H_0 be the height of the barometer corrected to 0° C., the density of mercury at 0° C. being ρ_0. Then the pressure in A $= H_0 g \rho_0 - H g \rho$, where H is the height indicated, g the intensity of gravity, and ρ is the density of mercury

at the steam temperature. Then, since $\frac{pV}{T}$ = constant for an ideal gas, V_0 is determined by the equation

$$\frac{V(H_0\rho_0 - H\rho)}{(273 + \theta_s)} = \frac{V_0 \times 76\rho_0}{273},$$

where θ_s is the steam temperature on the centigrade scale.

It ought to be noticed that the pressure in A cannot be found by subtracting H from H_0 since H represents a pressure due to a column of mercury at the steam temperature while the barometric height is measured directly at room temperature. To subtract these two heights is therefore absurd unless they are both corrected to the same standard temperature, viz. 0° C. The method adopted here is equivalent to this. Moreover, the volume of the tube will be found by calibrating it with mercury at room temperature so that a correction for the expansion of the glass has to be made in estimating the volume of vapour at the higher temperature. Another correction, although small, has to be made. It arises from the fact that when the volume of the tube is being estimated the curvature of the mercury is in the opposite direction from that in the actual experiment. A table giving the necessary correction will be found in Science Abstracts, A, 1910, No. 1563.

Dumas' Vapour Density Apparatus.—This apparatus, in addition to its historical importance as providing the first means of determining the vapour density of a substance, to-day furnishes us with the best means of determining such densities, when used by a skilled experimentalist, and full corrections are made for the rather numerous sources of error, which were not apparent to the earlier investigators. It is interesting to note that LORD RAYLEIGH used this method when he found that atmospheric nitrogen differed in density from that prepared chemically, a fact which ultimately led to the discovery of the inert gases argon, neon, krypton, xenon—gases which play an important part in modern atomic theory and also in industry.

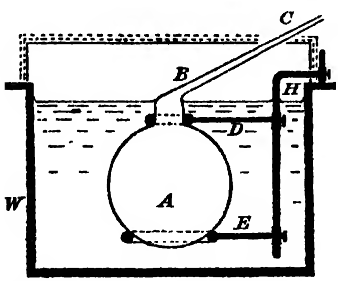


FIG. 12·13.—Dumas' Apparatus for determining the Vapour Density of a Substance.

The essentials of this method are as follows: A, Fig. 12·13, is a large glass globe, the neck of which is drawn out to a narrow and thin-walled tube, BC. This is supported by two metal rings, D and E, carried on a rod, H, attached to the outer casing of an iron bath, W, which may be filled with boiling water—or other sub-

stance. A copper lid, covered with asbestos, ensures that BC shall be at the temperature of the liquid boiling in W.

To carry out a determination of the vapour density of carbon tetrachloride (CCl_4), for example, the flask A is cleaned, dried, and its mass determined. Let m_1 be the mass in the scale pan when the flask is weighed in air, at temperature t and pressure p. Then

m_1 = apparent mass of bulb in air
= mass of bulb -- mass of air displaced by the material of the bulb at temperature t and pressure p.

About 5 cm.^{3} of carbon tetrachloride are then placed in the bulb, and when the water in W is boiling, the bulb and its contents are placed in position. The tetrachloride evaporates and displaces the air inside the bulb. Finally, when all the liquid has evaporated, the bulb remains filled with vapour at pressure p and temperature θ, the steam temperature under prevailing conditions. [A piece of polished metal held near to the jet C is no longer dimmed when vapour has ceased to issue from the bulb.] The tube BC is then closed, the bulb removed from the water, dried, and its total mass again determined. Suppose that m_2 is the mass in the scale pan when the balance is in equilibrium and the bulb filled with vapour. Let μ be the mass of the vapour in the bulb.

Then $m_2 = \mu +$ mass of bulb — mass of air displaced by the closed bulb at temperature t and pressure p.

Hence $m_2 - m_1 = \mu -$ mass of air required to fill the closed bulb at temperature t and pressure p.

To determine this mass of air, it is necessary to find the volume of the bulb. This is done by opening the neck of the bulb under water—the flask will completely fill with water if the experiment has been successfully carried out. Let V_1 be the volume of the flask at temperature t_1, as deduced from the mass of water it contains at temperature t_1. Let γ be the coefficient of cubical expansion of the material of the flask. Then

$$V_t = V_1[1 + \gamma(t - t_1)]$$

and

$$V_\theta = V_1[1 + \gamma(\theta - t_1)].$$

Let ρ_θ be the vapour density of the carbon tetrachloride at temperature θ and pressure p. Then

$$\rho_0 = \frac{273 + \theta}{273} \cdot \frac{76}{p} \cdot \rho_\theta.$$

Now $\mu = V_\theta \,.\, \rho_\theta$, so that ρ_0 may be deduced when μ is known.

Let σ_0 be the density of air under standard conditions, σ its density at temperature t and pressure p.

Then $$\sigma = \frac{273}{273 + t} \cdot \frac{p}{76} \cdot \sigma_0.$$

$\therefore$ mass of air required to fill the bulb at temperature t and pressure p

$$= \sigma V_t = \frac{273}{273 + t} \cdot \frac{p}{76} \cdot \sigma_0 \, . \, V_1[1 + \gamma(t - t_1)].$$

Hence ρ_0 may be deduced.

[In the above argument it has been assumed that the pressure and temperature of the air have remained constant throughout the experiment.]

Cooling Produced by Evaporation.—In consequence of the fact that it is the more rapidly moving molecules which escape from a liquid when evaporation occurs, it follows that the total kinetic energy of the molecules remaining behind must diminish, i.e. the liquid will be cooled. To maintain the rate of evaporation heat must be supplied—this we have already termed the latent heat of vaporization of the liquid.

Experiment.—Place a beaker containing ether on a wet piece of wood. By passing air from a bellows through the ether, cause this to evaporate rapidly. After a little while the beaker cannot be removed since the layer of water below will have frozen owing to the heat abstracted from it.

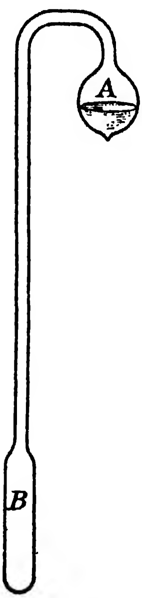


Fig. 12·14.—Wollaston's Cryophorus.

Wollaston's Cryophorus.—This instrument consists of two bulbs, A and B, Fig. 12·14, joined together by a fairly wide tube as shown. The bulb A contains water, the rest of the apparatus being filled with water vapour only. The bulb B is placed in a freezing mixture of ice and salt while the upper one is jacketed with a thick layer of cotton-wool to diminish the exchange of heat between A and its surroundings. The vapour in B condenses; it is replaced by vapour produced by the evaporation of water from A: more vapour condenses and more water evaporates. The necessary heat of vaporization is abstracted from the water itself so that its temperature falls, since heat can only pass very slowly through the cotton-wool to the water. After about fifteen minutes a layer of ice will have formed on the surface of the water and 'snow' will be visible in B.

EXAMPLES XII

1.—A quantity of air in contact with a liquid has a volume of 126 cm.^{3} at 19·2° C. under a pressure of 74·8 cm. of mercury. The pressure is increased to 141·8 cm. of mercury and the volume halved. If the temperature remains constant, calculate the saturation vapour pressure of the liquid at 19·2° C.

2.—How may the saturation vapour pressure of alcohol be measured between 50° C. and its boiling-point? An enclosed mass of air is saturated with water vapour at 100° C. On raising the temperature of the whole to 200° C. without change of volume the pressure increases to 2 atmospheres. Find, approximately, the pressure at 0° C.

3.—The space above the mercury in a barometer tube contains a little air, water vapour, and a drop of water. The length of the mercury column is found to be 73·5 cm. when the true height is 75·5 cm., and 74·6 cm. when the true height is 76·7 cm. Assuming that the top of the tube is 10 cm. above the mercury level in the first instance, calculate the pressure of the air and the vapour pressure of the water in the tube.

4.—Explain how the boiling-point of a liquid at temperatures somewhat above its normal boiling-point may be determined. A small liquid index encloses a volume of air in a uniform tube. If the length of the tube occupied by air is 20 cm. at 30° C., when the saturation vapour pressure of the liquid is 1·75 cm. of mercury, what will be the length when the temperature is 50° C., the saturation vapour pressure of the liquid then being 9·23 cm. of mercury? Height of barometer 76 cm.

5.—How would you study experimentally the relation between the saturation vapour pressure of water and the temperature, for temperatures above the normal boiling-point? A small glass bulb nearly filled with water is placed in an iron cylinder which is then heated in a vessel of boiling water. When the temperature is steady, the cylinder is hermetically sealed and the glass bulb broken by shaking. Discuss what will then be the value of the pressure inside the cylinder.

CHAPTER XIII

WATER IN THE ATMOSPHERE

Relative Humidity.—Our ideas concerning the conditions of the atmosphere, with reference to its moisture content, are often fallacious if they are formed without actual measurement. On a summer morning the presence of dew and slight haze shows that the air is saturated with water vapour. As the day progresses the heat of the sun warms the atmosphere and thereby enables it to carry more moisture without becoming saturated. Under such conditions it is often erroneously stated that the air is dry; actually it contains more moisture than before. It is only by comparing the masses of vapour present in a definite volume of air under various circumstances that the true facts can be ascertained. Instruments used for this purpose are called ***hygrometers,*** whilst the ratio of the actual amount of water present to that required to saturate it at the same temperature is referred to as the ***relative humidity,*** or ***humidity,*** of the atmosphere.

If m is the actual mass of vapour present in a given volume of air, and M the mass required to saturate it at the same temperature, the relative humidity is $\frac{m}{M}$. But if the vapour obeys Boyle's law, which it does approximately, this same ratio is equal to $\frac{p}{P}$, where p is the partial pressure of the water vapour present and P is the saturation vapour pressure of water at a temperature equal to that of the air.

The Dew-point.—When moist air is cooled, a temperature is soon reached when the quantity of moisture present is sufficient to saturate the air; any further cooling causes some of the vapour to be deposited on surrounding objects in the form of dew. If the 'surrounding objects' are not visible the deposition takes place on small nuclei—dust, etc.—and fog is produced. The first hygrometers invented were designed to estimate the dew-point, for then the partial pressure of the water vapour is known since this is equal to the saturation vapour pressure of water at the temperature of the

dew-point. The vapour pressure of water at the existing tempera ture may be obtained from tables and the relative humidity may then be calculated. The essential features of a good hygrometer are that an observer should be able to ascertain the exact instant when dew begins to be deposited, and he should also be able to know the temperature of the surface on which the deposition occurs.

Daniell's Hygrometer.—This consists of two bulbs joined together as in a cryophorus, the enclosed liquid being ether. The lower bulb contains the liquid while the upper one is covered with a muslin bag on to which ether is poured. The rapid evaporation of the ether lowers the temperature here so that some vapour inside this bulb condenses. This causes the ether in the lower bulb to evaporate, a cooling effect being noticed. As the cooling proceeds dew begins to be deposited on a gold band on the outside of the bulb. The temperature is recorded by a thermometer inside the bulb where it is undoubtedly lower than that of the gold band. To overcome this difficulty the temperature is observed at which the dew disappears when the temperature rises again ; the mean is taken as the dew-point. The upper bulb must only contain ether vapour—a little liquid ether may be allowed and some liquid certainly collects in this bulb while the instrument is in use. The mass of liquid in this bulb must never be allowed to become large, for if it does the rate of evaporation of the liquid in the lower bulb is seldom rapid enough for dew to be deposited. Although Daniell's hygrometer is the oldest form of such instrument it is unfortunately the most objectionable. Amongst the several objections we may mention that it is made of glass, a poor conductor of heat, and therefore does not assist in the establishment of a uniform temperature. In addition, although a mean value of the temperature is observed it cannot be the true dew-point since the outside of the instrument is always hotter than the inside. Moreover, the air around is filled with ether vapour, and the liquid in the hygrometer bulb is not stirred. Many of these objections do not apply to Regnault's hygrometer.

Regnault's Hygrometer.—This apparatus consists of a glass tube A, Fig. 13·1, the lower end of which is attached to a highly polished thin silver thimble. The tube contains ether, and a piece of quill tubing passes through a cork, nearly to the bottom of A. A standardized thermometer also passes through the cork, its bulb being in the ether. A side tube is connected, through the tubular stand, to an aspirator, C, preferably placed at a considerable distance away so that the moisture content of the atmosphere shall not be disturbed by it. The withdrawal of air from the aspirator causes air to bubble through the ether and this produces a rapid evaporation with a consequent cooling. The process is continued until

moisture is deposited on the thimble: the temperature is recorded and the flow of air reduced so that one bubble of air passes in 5 or 10 sec. The cooling now produced is insufficient to maintain the temperature low enough for the moisture to remain on the thimble, but the bubbles are necessary in order that the liquid may be thoroughly

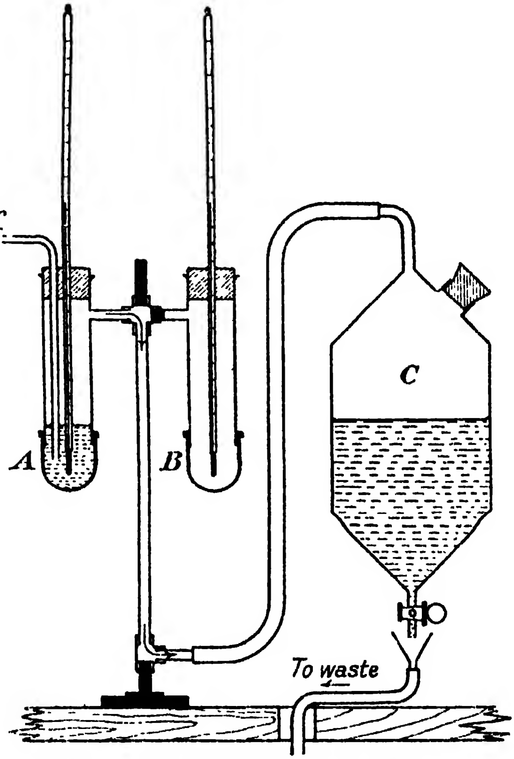


FIG. 13·1.—Regnault's Hygrometer.

stirred. The temperature at which the moisture disappears is observed, and the mean temperature taken as the temperature at which the moisture content of the atmosphere is sufficient to saturate the air, i.e. it is the ***dew-point***. The bulb B does not contain ether and is simply used as a comparator; both tubes are protected from the operator by means of a large sheet of glass.

Dines' Hygrometer.—A reservoir, A, Fig. 13·2, is filled with water cooled by ice. A second chamber, C, communicates with A through a long tube. E is an exit tube. The flow of water is controlled by a tap, T. The upper part of C is closed by a piece of black glass or a silver plate. The thermometer D records the temperature when dew appears. T is then closed and the temperature at which

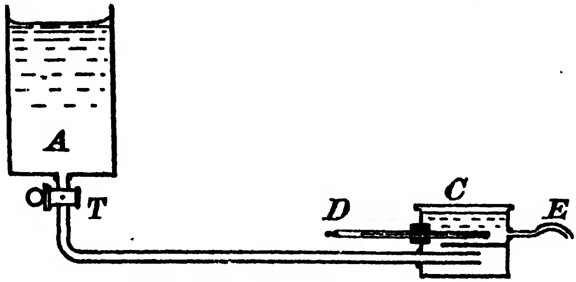


FIG. 13·2.—Dines' Hygrometer.

the dew disappears noted. The mean of these two observations is the dew-point but, as in Daniell's instrument, it is not a reliable estimate of the dew-point.

The Wet and Dry Bulb Hygrometer.—For field observations it is inconvenient to use a Regnault hygrometer; instead, use is made of MASON's wet and dry bulb hygrometer. Two thermometers are placed side by side, the bulb of one being surrounded by muslin, kept moist by means of a piece of cotton wick dipping into a small vessel of water. The 'wet' bulb indicates a lower temperature than the dry bulb on account of the heat absorbed during the constant evaporation of the water which occurs there. This difference depends upon the relative humidity of the atmosphere; to enable this to be calculated tables have been prepared showing the relative humidity corresponding to this temperature difference under various conditions.

Instead of using tables in connexion with this hygrometer, the following formula is sometimes used to calculate the pressure, p, of the water present in the air. If t is the air temperature, t_w that of the wet bulb, p_w the saturation vapour pressure of water at the temperature of the wet bulb, P the atmospheric pressure, then

$$p = p_w - AP(t - t_w)$$

where A is a numerical factor. Attempts have been made to establish the formula theoretically, but they are not satisfactory.

Wet and Dry Bulb Hygrometers of the Ventilated Type: Psychrometers.—The factor A in the above equation differs materially according to whether or not the wet bulb is in quiet or in moving air. In view of this the instrument has been regarded as a notoriously unreliable one. By modifying the construction, however, so that air at a definite velocity is drawn past the thermometer bulbs, it may be converted into a satisfactory instrument. 'This important fact was demonstrated by the Italian physicist BELLI in 1830, and in view of the simplicity of the device it is somewhat surprising that the stationary form of wet and dry bulb hygro-

meter is tolerated at all to-day,' says Griffiths in a Discussion on Hygrometry before the Physical Society. He showed that if the velocity of the air exceeds about three metres per second the factor A assumes a constant value for a particular instrument. It is determined from simultaneous readings obtained with this hygrometer and with Regnault's [modern type, cf. p. 268].

To obtain a sufficient velocity of air past the thermometers they are secured to a rod and whirled round, or an electric motor may be used to draw air past them. The following instrument designed by GRIFFITHS and an improvement on an earlier pattern by ASSMANN is known as a ***tubular psychrometer.*** It consists of a steel tube, Fig. 13·3 (*a*), in which the thermometers are placed. By means of a fan coupled to an electric motor, air is drawn past the bulbs of the thermometers. After the sack round the wet bulb has been moistened the instrument may be used for 40 minutes without replenishing the supply of water. A glass window is inserted in the tube so that the thermometers may be read. An advantage of this instrument is that it is not necessary for the observer to be in the room where the humidity is being determined.

The Chemical or Gravimetric Hygrometer.—In this method air at a known mean temperature, indicated by thermometers placed in the immediate vicinity of the apparatus, is drawn over pieces of pumice soaked with concentrated sulphuric acid (this is more efficient than calcium chloride) and contained in tubes D, Fig. 13·4. The pieces of pumice stone must not be too small or the amount of acid excessive, so that air passes freely through the tubes, the pressure of the air in A then being atmospheric. The aspirator A contains water, the vapour of which is prevented from reaching the absorption tubes by means of calcium chloride contained in a bottle B. The cork in the aspirator supports two glass tubes bent at right angles; one acts as a siphon and the other serves to connect the aspirator with the rest of the apparatus. The assembling of the apparatus is facilitated by using absorption tubes of the type shown in the diagram. To prevent air leaking into the apparatus the corks are pushed well into the tubes and the necks of the bottles A and B, and then covered with molten wax[1] which is allowed to solidify. The tube C contains asbestos wool which serves to prevent dust particles from reaching the absorption tubes.

Before using this hygrometer it is essential to discover whether or not the apparatus is leaking. To do this, the stop-cock T_1 is closed, while the stop-cocks T_2 and T_3 are opened. On opening the stop-cock T_4 attached to the delivery tube no water should exude from the apparatus, or rather it should not continue to

[1] A suitable wax consists of a mixture of beeswax and vaseline—about equal parts being melted together and then allowed to solidify.

do so—it will flow out slowly at first until pressure conditions inside the apparatus have adjusted themselves to correspond to the difference in level between the water in A and that of the exit of the delivery tube. If water continues to flow out there is a leak, and this must be rectified before proceeding further. It is essential that all the air finding its way into the aspirator should have passed through the drying tubes.

T_4 is then closed and the drying tubes removed in order that their mass may be ascertained. During the above process the pressure inside A will have become atmospheric. All taps are then

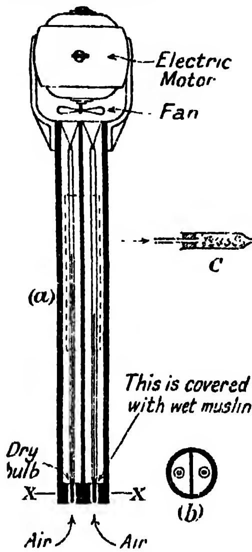


FIG. 13·3.—(*a*) Wet and Dry Bulb Hygrometer of the Ventilated Type (a psychrometer).
(*b*) Section across XX.

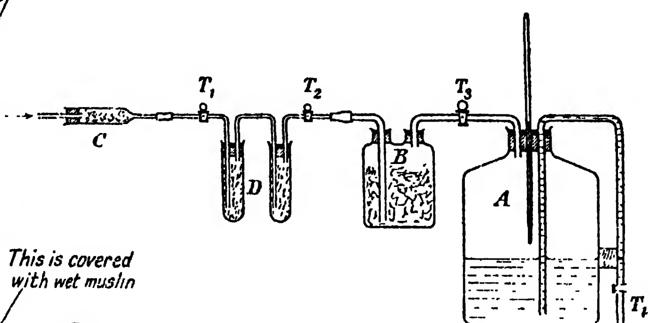


FIG. 13·4.—A Chemical or Absorption Hygrometer.

opened, T_4 being adjusted so that the rate of flow of the water is about one litre per minute. After about half an hour, T_4 is closed and the increase in mass of the drying tubes determined. The mass of the water which has escaped from the apparatus is also determined.

If all corrections are neglected, the mass of water per cubic metre of the air is then deduced: if the amount required to saturate an equal volume of air at the same temperature and pressure is known, the relative humidity may be calculated. The objection to this method is that it occupies a considerable time and only gives a mean value of the humidity.

Theory of the Gravimetric Hygrometer.—For the sake of simplicity we shall assume that θ is the temperature of the air and the aspirator, and that it remains constant. Let V be the volume (in cubic metres) of water run out from the aspirator. Let P be

the barometric pressure; p the actual pressure of the water present in the air and p_s the saturation vapour pressure of water at $\theta°$ C. Since the dry air leaving the drying tubes and passing into the aspirator becomes saturated with water vapour the total pressure P inside the aspirator is made up of p_s, the saturation vapour pressure of water at $\theta°$ C., and $(P - p_s)$, the partial pressure of the dry air.

But this air was moist when it entered the hygrometer, its partial pressure then being $(P - p)$. Then V_1, the volume of air entering the apparatus, is, by Boyle's law, given by

$$V_1(P - p) = V(P - p_s).$$

Let μ be the increase in the mass of the drying tubes after an experiment. Then the mass of vapour present per cubic metre is

$$\frac{\mu}{V_1} = \frac{\mu}{V} \cdot \frac{P - p}{P - p_s}.$$

Now, unfortunately, in the above equation p is not known. It may be calculated as follows. Let σ be the relative density of water vapour with respect to air at the same temperature and pressure—then under the low pressures here contemplated σ is a constant. If M is the mass of a cubic metre of air at S.T.P., the mass of a cubic metre of air at pressure p and temperature θ is

$$\frac{Mp}{76(1 + \alpha\theta)}. \quad [\alpha = \tfrac{1}{273} \text{ deg.}^{-1} \text{ C.}]$$

The mass of a cubic metre of water vapour at pressure p and temperature θ is therefore

$$\frac{Mp\sigma}{76(1 + \alpha\theta)} = \frac{\mu}{V} \cdot \frac{P - p}{P - p_s}.$$

From this equation p is determined. The relative humidity is then given by $\left(\frac{p}{p_s} \times 100\right)$ per cent.

[The mass of water present in a cubic metre of air measured under existing conditions may then be calculated—it is not given at once by this method as ordinarily supposed. The above quantity is termed the ***absolute humidity*** of the atmosphere.]

A Modern Form of Regnault's Apparatus.—The hygrometer itself, consisting of a silver thimble at the end of a glass tube, is mounted in the lid of a wooden box one half of which may be moved about hinges attached along that edge of the box passing through A, Fig. 13·5. Air is forced through some ether placed in the thimble by gently squeezing a rubber ball attached to the air inlet by rubber tubing. The surface of the thimble is viewed through a double glass window, the interior of the box being illuminated by a lamp L placed outside the box. Before a determination of the humidity

at any station is made, the lower half of the box is made to oscillate several times to ensure that the air inside the box is an average sample. The box is then closed and the experiment conducted in the usual way. The particular advantage of this apparatus is that the hygrometer is screened from the observer while observations are being made. This is very essential when the temperature is low, for it must be remembered that at 0° C. less than 5 gm. of water is sufficient to saturate a cubic metre of air.

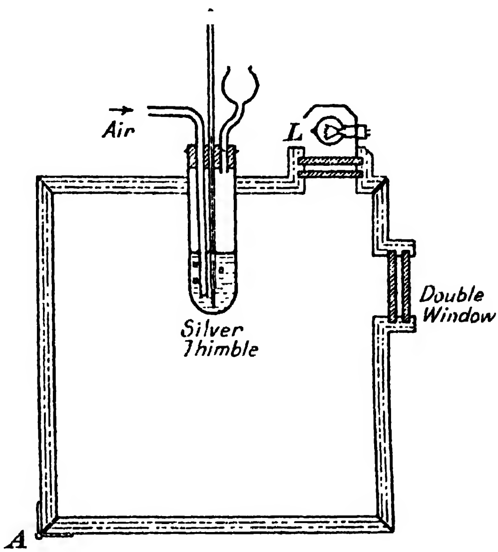


FIG. 13·5.—N.P.L. Form of Regnault's Hygrometer.

Cloud and Mist.—The production of *mist* or *cloud* is one of the results produced by the condensation of moisture in the air—a cloud simply being mist at a high altitude. In order for such small drops of water to be formed some dust particles or electric charges must be present upon which the water vapour may condense. The drops must be extremely small since they do not fall to the earth—the viscous nature of the atmosphere is slight, but nevertheless sufficient to prevent any rapid motion of these particles. There is a tendency, however, for these drops to coalesce, and when they are sufficiently large, *rain* is precipitated.

Snow and Hail.—Snow is probably a consequence of the direct passage of water vapour into the solid state. Hail is most likely due to the freezing which takes place when rain-drops pass through strata of air where the prevailing temperature is below 0° C. By cutting a hailstone in halves it has been shown that such a stone may consist of several distinct layers, proving that moisture has condensed upon the original piece of hail several times, and, after each condensation, freezing has occurred.

Dew and Frost.—The small drops of water which are seen clinging to stones, etc., in the early hours of a summer morning and at other times are referred to as *dew*. These drops have not been

produced in the regions remote from the surface of the earth, but in the immediate vicinity of the earth's surface. Dew is generally noticed when clouds have been absent. The absence of cloud permits heat to be radiated into space and this loss of heat is followed by a local lowering of the temperature. This lowering is much more marked when stones, etc., are present, for these are good radiators, and if the temperature is lowered below the dew-point the appearance of water-drops on the cooled object is the result.

When there is no wind the layers of air near to the objects from which heat is being radiated are cooled more rapidly, so that the production of dew is favoured. The conditions favourable for the formation of dew were first stated by Wells (1812) in a celebrated 'Essay on Dew,' and are as follows:

(i) there should be a cloudless sky,

(ii) there should be no wind,

(iii) the relative humidity of the atmosphere should be high.

In 1886 Aitken extended the above theory. He maintained that there are two types of dew:

(i) that which depends on the water vapour present in the atmosphere,

(ii) that which results from the water given off by the leaves of plants. This is emitted as a vapour and under normal conditions passes at once into the unsaturated air. When the air near the leaves is saturated with water vapour, then that which comes from the leaves appears as dew on them.

If the dew-point is below 0° C., and the temperature still lower the water is deposited as ***hoar-frost***. When the temperature is below 0° C. but the air not saturated with moisture, then the prevailing conditions are known as a ***black-frost***.

EXAMPLES XIII

1.—Write a short essay on the measurement of the humidity of the atmosphere.—(L. '28.)

2.—Define the terms *relative humidity*, *absolute humidity*, and *saturated vapour*. Describe and discuss the method due to Regnault and the chemical (or gravimetric) method of determining the relative humidity of the atmosphere.

3.—Explain how the *absolute humidity* and the *relative humidity* of the air may be measured.

In certain conditions of weather, the walls and tiled floors in a house may become very damp. How do you account for this? How may it be prevented?

CHAPTER XIV

THE DYNAMICAL THEORY OF HEAT

Early Theories of Heat.—From the early days of science down to the beginning of the nineteenth century two rival schools had expressed their opinions concerning the nature of heat. The one regarded heat as a subtle fluid which permeated the pores of a body; the other maintained that heat was due to the motions of the molecules. Neither theory was well founded—in fact, we may use them to compare the way in which research was prosecuted by the Ancient Greeks and that adopted to-day—or rather, the ways were always the same, only the various factors were assessed differently. The philosophers of the Classical Era were satisfied with very few facts and proceeded at once to form a theory when they had become cognizant of them. Nowadays, it is not until many facts from a multitude of various sources have been obtained that a theory is attempted.

The Caloric Theory.—This theory attributed heat to the presence of a self-repellent and all-pervading fluid. It was attracted by all forms of ordinary matter and an increase in temperature was due to a gain in caloric, the resulting expansion being due to the mutual repulsion of its particles. It was generally held to be imponderable. The conduction of heat was attributed to the flow of caloric from a higher to a lower temperature—the driving agent being the self-repellent nature of this imponderable fluid.

Hard Facts for the Calorists.—The adherents of the caloric theory were well acquainted with the fact that heat may be produced by friction as when a savage rubs two pieces of dry wood together to kindle his fire, or when a grinding wheel wears away the surface of a metal it is polishing, and the heat developed is so great that the abraded material is raised to incandescence. The calorists explained this by making the arbitrary but incorrect assumption that the thermal capacity of a substance was less in the powder form than otherwise. But still harder facts were in store for them. In 1798, while engaged in the boring of cannon at Munich, COUNT RUMFORD noticed the large amount of heat developed in this process. To test the matter further he used a blunt borer. The heat was still

generated, although the amount of abraded metal was negligibly small. The calorists, however, held steadfastly to the tenets of their theory, but when DAVY showed that ice could be melted by rubbing two pieces together the death-knell to this theory was sounded (1799). According to the calorists, the friction had caused caloric [heat] to be squeezed out from the ice, i.e. the thermal capacity of water would be less than that of ice. This was contrary to experiment. Davy therefore concluded that this imponderable fluid called caloric did not exist, but that the motion of the ice molecules was increased by the rubbing, and that this increased motion revealed itself in the melting of the ice.

Joule's Experiments and the First Law of Thermodynamics.—During the years 1842–1848 JOULE [of Manchester] made some classical experiments on the relationship between mechanical work and heat. He showed that, irrespective of how the work was done, the heat generated was directly proportional to the work done. If W is the quantity of work done [usually measured in ergs], Q the amount of heat (calories) produced, the above statement is expressed by the equation

$$W = JQ,$$

where J is a constant, known as ***the mechanical equivalent of heat.***

Modern work has shown that $J = 4{\cdot}184 \times 10^7$ erg. cal.$^{-1}$, i.e. $4{\cdot}184 \times 10^7$ ergs of work must be developed to produce one calorie. In the f.p.s. system of units 1440 ft. lb.-wt. of work are necessary to raise the temperature of 1 lb. of water 1° C.; this is equivalent to 772 ft. lb.-wt. deg.$^{-1}$ F.

Although Joule was the first to make any accurate measurements on the relationship between heat and work, the fact that there might be a connection was suspected in 1842 by a German physician, MAYER. He noticed that the blood in the *veins* was brighter in colour for persons in tropical lands than elsewhere. He argued that the gain in heat by a person in the tropics was greater than that in cooler regions; the decomposition of the blood was therefore less severe and so it had a brighter colour.

A diagrammatic sketch of Joule's first apparatus is shown in Fig. 14·1 (*a*). The water in a calorimeter A was churned by a paddle carrying eight vanes, the mere rotation of the water being prevented by four fixed vanes inside the calorimeter—cf. Fig. 14·1 (*b*). The calorimeter was placed inside a constant temperature enclosure and supported on three ivory feet. The box-wood cylinder C acted as a heat insulator between the calorimeter and the axis of the paddle. The motion was imparted to the paddle by means of two large masses, M_1 and M_2, capable of descending through a fixed distance. These masses were carried by strings

fastened to the axles of two large pulleys, BB, mounted on 'friction wheels.'

If m_1 and m_2 are the masses of the large metallic blocks, and h the height of fall, the work done in one descent is $(m_1 + m_2)gh$, for the force $(m_1 + m_2)g$, i.e. the total weight, acts through a distance h. The blocks were allowed to fall n times to produce an appreciable rise in temperature of the calorimeter and its contents, so that the total work done was $n[(m_1 + m_2)gh] = W$. The wooden cylinder, D, round which the string passed, was attached to the axis of the paddle by a taper pin, P, and this was removed to prevent the paddle rotating while the masses were being raised.

If C is the thermal capacity of the calorimeter and its contents,

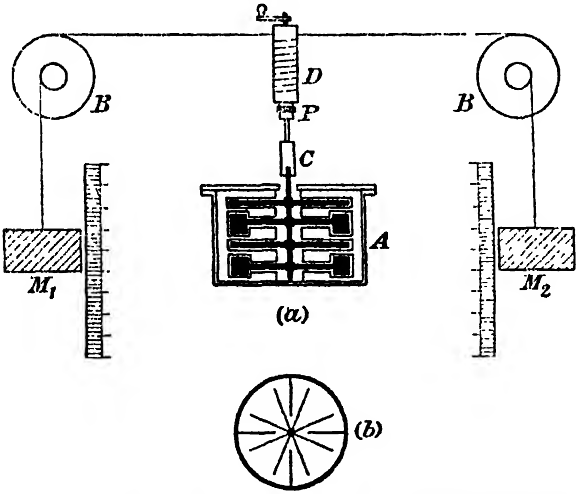


FIG. 14·1.—Joule's Original Apparatus for Determining J.

θ the observed rise in temperature, the quantity of heat produced, measured in calories, is $C\theta = Q$, so that

$$J = \frac{W}{Q} = \frac{n[m_1 + m_2]gh}{C\theta}.$$

In deriving this formula we have assumed that there has been no exchange of heat between the calorimeter and its surroundings, and that the whole of the energy possessed by the blocks has been given to the calorimeter and its contents. Joule made a correction for the heat exchange by observing the rate of cooling of the calorimeter. Another correction is also necessary, for when the blocks reached their lowest position they were moving with a certain definite velocity—say v—so that their kinetic energy was $\frac{1}{2}(m_1 + m_2)v^2$. This must be subtracted from the total potential

energy to obtain the energy imparted to the calorimeter. A further correction is necessary for the work spent in overcoming friction in the moving parts outside the calorimeter. To estimate this Joule disconnected the drum D from the paddle at C and the cord from the pulleys was passed round it in such a manner that as one block fell the other rose. To produce this motion a small mass μ was placed on one of the masses, it being adjusted so that the motion was uniform. The resistance due to friction was therefore μg dynes and the total work spent in overcoming it $nh\mu g$. The energy actually given to the water before any had been lost by radiation, etc., was therefore

$$n[(m_1 + m_2)(gh - \tfrac{1}{2}v^2) - \mu gh],$$

so that the correct equation is

$$J = \frac{n[(m_1 + m_2)(gh - \tfrac{1}{2}v^2) - \mu gh]}{C\theta_1},$$

where θ_1 represents the corrected rise in temperature, i.e. the temperature rise obtaining in the absence of heat losses.

Joule's Second Apparatus.—About thirty years after the above experiments had been completed the British Association requested Joule to repeat his work because, in the meantime, some experiments in which electrical energy had been converted into heat had been made and there was a discrepancy of 1 per cent. between the values of J obtained by the two methods. In the earlier form of apparatus the rise in temperature was only 0·5°, and the heat was not generated continuously. In addition, the paddles did not experience a steady resistance, for it was a maximum when the paddles moved through the openings in the fixed vanes. Originally there had been four vanes and eight paddles so that these maxima occurred eight times per revolution. In the second apparatus, Fig. 14·2 (*a*), there were four vanes and two sets of paddles each with five arms as at (*b*). Consequently there was never more than one paddle passing through an opening in the vanes at the same instant, yet such an event occurred forty times in each revolution. In consequence of this the driving torque was more steady and there was less vibration set up in the apparatus.

The calorimeter, A, was supported so as to be free to rotate with the paddle, but such a rotation was prevented by applying a couple. For this purpose a fine silk cord passed round a groove on the surface of the calorimeter. This cord passed over two pulleys, BB, and carried scale pans, SS, suitably loaded. The motion was obtained by means of the hand-wheels, CC, a heavy flywheel, D, being attached to the vertical shaft to assist in steadying the motion. A conical bearing, E, supported the vertical shaft. After some preliminary work Joule found that the friction of the bearings was

not constant and so invented the hydraulic supporter, FG, to reduce the pressure on the bearing. This consisted of two co-axial vessels, F and G, the space between them being filled with such a quantity of water that the three uprights attached to the lid of the hollow vessel and in contact with the base of the calorimeter just relieved the thrust on the bearing.

In making an experiment the calorimeter was filled with a known mass of water and its temperature noted. The thermometer was

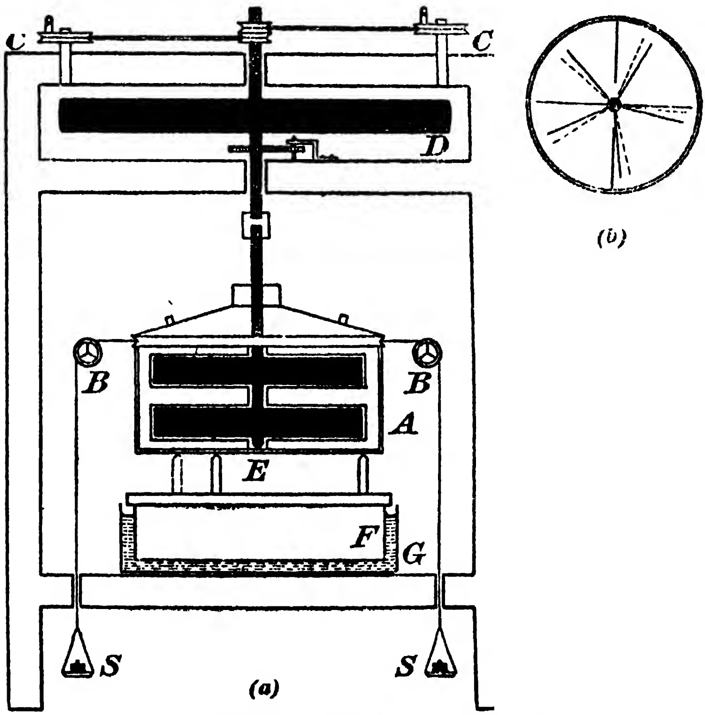


FIG. 14·2.—Joule's Second Apparatus.

removed and the wheels rotated so that the two masses were in equilibrium at a distance of one foot from the floor. They were maintained in this position for thirty-five minutes, after which the paddles were brought to rest and the final temperature recorded. The number of revolutions was counted mechanically. If $(m_1 + m_2)$ is the sum of the two masses suspended from the silk cord, and r is the radius of the calorimeter, the work done in n revolutions is $2\pi n(m_1 + m_2)gr$. This is equal to $JC\theta$, where C is the thermal capacity of the calorimeter and its contents, and θ its rise in temperature. From these observations J was calculated.

At a later date ROWLAND made an elaborate series of experiments using a calorimeter and method similar to the above. The calorimeter was rotated mechanically to obtain a large rise in temperature in a relatively short time. This diminished the correction for heat lost by radiation, etc.—in fact, the correction was only one-fiftieth of that in Joule's experiment for the same rise in temperature.

Laboratory Method for the Determination of J.—The instrument used, which is based upon the original design by CALLENDAR, is illustrated in Fig. 14·3 (*a*). The principle of operation is as follows :— Mechanical energy is dissipated by means of a special brake rubbing on the outside of a rotating brass drum, or calorimeter, containing water, and the heat energy developed is deduced from observations on the rise in temperature of the water and the water equivalent of the calorimeter and its contents. The calorimeter drum, A, is rotated

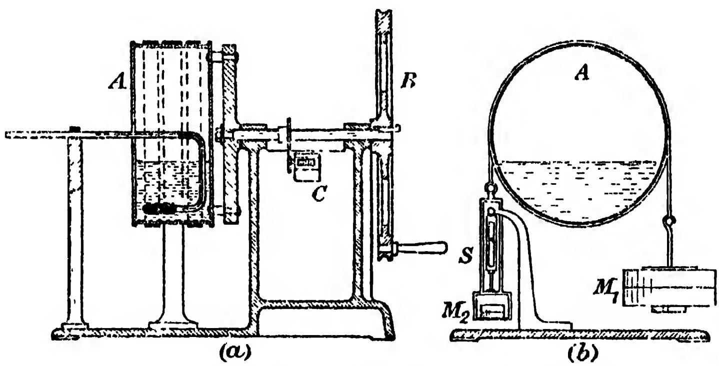


FIG. 14·3.—Callendar's Apparatus for J (Mechanical).

about a horizontal axis by means of a driving wheel, B. This wheel can be driven either by hand or by a 0·1 h.p. motor through suitable reduction gearing. The number of revolutions made by the calorimeter drum is automatically recorded by the counter C. The brake consists of a silk belt arranged to form $1\frac{1}{2}$ complete turns round the drum. Unequal and adjustable masses, M_1 and M_2, are suspended from the ends of this belt, and are automatically maintained in a position of floating equilibrium by a light spring balance S acting in opposition to the weight of M_2—see Fig. 14·3 (*b*). The extreme flexibility of the belt ensures that the difference between the weights of the loads at the two ends of the belt measures the friction. The rise in temperature of the water in the calorimeter is observed by means of a mercury thermometer inserted through an axial opening in the front of the cylinder. The bulb of this thermometer is bent round so that it is totally immersed in the

water. To ensure good results the surface of the drum should be kept as smooth and bright as possible and the belt clean and dry.

The work done, W, is the product of the difference of the *weights* on the two sides multiplied by the number of revolutions and the circumference of the drum [cf. p. 67]. Thus if m_1 and m_2 are the masses of the loads respectively in grams, s the reading of the spring balance in gm.-wt., then the weight on one side is m_1g, while the effective weight on the other is $(m_2 - s)g$, since the spring balance acts in opposition to the weight of M_2. Hence

$$W = 2\pi nr[m_1 - (m_2 - s)]g \text{ erg.},$$

where r is the radius of the drum, and n the number of revolutions. If C is the thermal capacity of the calorimeter and its contents, θ the rise in temperature, then $Q = C\theta$, so that

$$J = \frac{2\pi nr[m_1 - (m_2 - s)]g}{C\theta} \text{ erg. cal.}^{-1}$$

To apply a correction for the heat lost, the initial temperature of the water should be two or three degrees below that of the room. The drum is set in motion and the instant when the temperature is that of the room noted. The duration of the experiment from this stage is noted, and the calorimeter then left to cool for the same time —the drum should be rotated but the band removed so that the heat is lost under the same conditions as in the actual experiment. Let $\Delta\theta$ be the fall in temperature during this interval. Since the mean excess of temperature over the surroundings is half the final excess, the rate of loss of heat at the end of the experiment is twice the mean rate of loss of heat during the actual experiment. If $\frac{1}{2}\Delta\theta$ is added to θ, a correction for the heat lost from the calorimeter will be made. Hence

$$J = \frac{2\pi nr[m_1 - (m_2 - s)]g}{M(\theta + \frac{1}{2}\Delta\theta)}.$$

In an actual experiment $n = 662$, $2r = 14{\cdot}9$ cm., $m_1 = 4265$ gm., $m_2 = 202$ gm., $s = 37$ gm.-wt. (mean value), $M = [300 + (384 \times 0{\cdot}092)]$ gm., $\theta = 8{\cdot}50°$ C., $\Delta\theta = 0{\cdot}80°$ C. [Time = 5 mins.] Hence

$$J[300 + (384 \times 0{\cdot}092)]\,[8{\cdot}5 + 0{\cdot}4] = 3{\cdot}14 \times 14{\cdot}9 \times 662 \times 4100 \times 981$$

$$\therefore J = 4{\cdot}18 \times 10^7 \text{ erg. cal.}^{-1}.$$

Callendar's Electrical Method of Determining J.—We have already described the continuous-flow method of determining specific heats. In that method the value of J was assumed: we may now reverse the process and, having defined the mean specific heat of water over a range in temperature from 15° C. to 20° C. to be 1 cal. gm.$^{-1}$ deg.$^{-1}$ C., proceed to calculate J. Hitherto only

a laboratory form of apparatus has been described, but, on account of the importance of this type of calorimetry both in industry and pure science at the present time, the actual calorimeter will now be described.

The electrical method is very advantageous, for the supply of heat can be controlled accurately. The method was suggested at an early date and even JOULE made measurements in this way, but, at that time the electrical units were not known accurately. Joule did not place much reliance upon his results, and the fact that there was a discrepancy of 1·5 per cent. between the electrical and mechanical methods induced Rowland to perform the experiments already mentioned. JAMIN determined J by an electrical method and detected a variation in the specific heat of

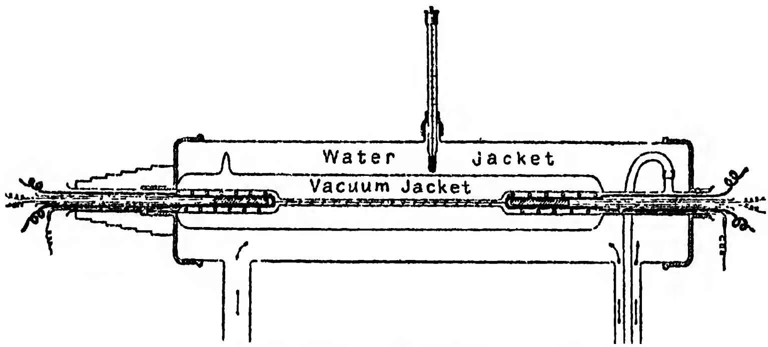


FIG. 14·4.—Callendar's Apparatus for J (Electrical).

water: the variation he obtained was twenty times that discovered by Regnault.

GRIFFITHS, at a later date, employed an electrical method, and diminished the external loss of heat by enclosing the calorimeter in a vacuum, but the vacuum he obtained was not exceptionally efficient for this purpose.

SCHUSTER and GANNON also used an electrical method, but they used mercury thermometers so that their results are open to the objections raised against such thermometers.

In all the above experiments the thermal capacity of the calorimeter was not always even small; Callendar's steady flow calorimeter, Fig. 14·4, is so used that the thermal capacity does not appear at all in the calculations. A steady stream of water passing through a fine tube is heated by an electric current through a central conductor. The increase in temperature between inflow and outflow was determined by a pair of platinum thermometers arranged differentially. This enabled the temperature difference to be

measured with an accuracy unattainable with mercury thermometers. To minimize the external loss of heat the flow-tube and thermometer pockets were sealed in a vacuum jacket. The whole was surrounded by a water jacket at a temperature of the inflow so that the exchange of heat between the calorimeter and its surroundings occurred under constant conditions. The current and potential difference along the conductor were measured by a potentiometer method.

If these are I amperes and V volts respectively, the energy dissipated per second is VI $\times$ 10^7 ergs. If m is the mass of water flowing per second and θ the observed rise in temperature the heat produced per second is $m\theta$ calories. Hence

$$J = \frac{VI \times 10^7}{m\theta} \text{ erg. cal.}^{-1}$$

In this equation we have assumed that the heat lost is zero. The actual method of determining this heat loss has already been explained [cf. p. 206]: only there we assumed J to find the specific heat of a liquid, whereas now we assume the specific heat of water to be unity and determine J. The specific heat of water varies with the temperature: it is taken to be 1 cal. gm.$^{-1}$ deg.$^{-1}$ C. over the range 15° C.–20° C., so that when J is being determined the experiment must be made over this range of temperature. The great merit of the continuous flow calorimeter is that when once J is known the capacity of a liquid for heat over a small temperature interval may be determined accurately. By surrounding the calorimeter with a water jacket at the temperature of the inflow at these higher temperatures, a further advantage was gained—the actual heat lost was made smaller than if the jacket had been at room temperature. We have also seen how the method was applied to measure the specific heat of a gas at constant pressure.

A Simple Hot Air Engine.—In the experimental determinations of the mechanical equivalent of heat described above, a measured amount of mechanical or electrical energy has been dissipated and the corresponding amount of heat determined calorimetrically. A device whereby heat is converted into mechanical energy is termed a heat engine. A simple but nevertheless interesting form of heat engine may be constructed as follows: A, Fig. 14·5, is a silica flask connected to a U-tube of the dimensions shown and containing mercury. D is a stop-cock. The air in A is heated by means of a bunsen burner. The pressure inside A increases so that the mercury is pushed down in one limb of the U-tube and raised in the other. D is opened for a fraction of a second and then closed. The mercury returns to the equilibrium position, but its momentum causes it to overshoot this mark: if

the amount of mercury in the U-tube has been properly adjusted it will begin to oscillate and continue to do so as long as heat is supplied to the air in A.

The operations which occur are as follows: heat is communicated to the air so that the pressure inside the apparatus increases. When the mercury descends in B and rises in C, the pressure of the air in A is reduced and a certain amount of air finds its way into B. Here it loses heat to the cooler portions of the apparatus, so that the pressure of the air is reduced. The air which had been driven from A returns, and is then heated once more: the process continues if the period of oscillation of the mercury in the U-tube is correct with regard to the rate at which heat is taken in and given out by the air in A.

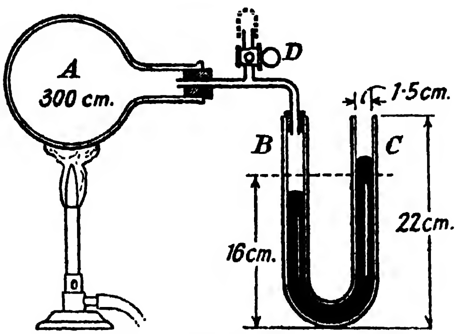


FIG. 14·5.—A simple Hot-Air Engine.

The characteristic features of a heat engine and those found in the above are: (i) the working substance (air) which expands and does work when thermal energy is supplied to it, (ii) the source of heat, (iii) the sink or cooling arrangement whereby the working substance is cooled after it has performed work and returned to its original state in the engine so that it may again take part in the working of the engine.

Isothermal and Adiabatic Changes.—Graphical Representation of the State of a Substance.—The pressure and volume of a gas may be indicated by co-ordinates: if either or both vary, the point representing the state of the gas will trace out a curve. This method of representing the state of a substance was devised by WATT for the purpose of calculating the work done by a steam engine. In general, as the state of the gas varies the temperature will also change, i.e. there must be an exchange of thermal energy between the gas and surrounding objects in order that its state may change. It may happen that the temperature remains constant although the state of the gas varies—such a change is said to be ***isothermal,*** and the curve representing the relation between p and v under these conditions is known as an ***isothermal curve.*** If a gas changes its state in such a way that there is no transfer of heat between it and its surroundings, the change is termed ***adiabatic***; the curve representing the relation between

p and v when no heat is supplied to or removed from the gas is known as an ***adiabatic.***

Let us examine the two changes more closely. Suppose that the gas is contained in a cylinder fitted with a piston. Let us assume that the gas is slowly compressed. According to the kinetic theory of gases, the temperature of a gas is determined by the mean kinetic energy of its molecules. When the piston is moved inwards work is done on the gas, i.e. energy is imparted to the gas molecules so that the temperature of the gas rises. If the rate at which work is done on the gas is very slow, there will be an exchange of heat between the gas and the walls of the cylinder, and when the above rate is infinitely slow the average kinetic energy of the gas will remain constant. Such is an ***isothermal change.***

When the piston is pushed in very rapidly, however, there is no exchange of heat between the gas and the walls, so that the temperature of the gas rises, and although the temperature has increased, no heat has been imparted to or abstracted from the gas, i.e. the change is ***adiabatic.***

If, in addition, at all stages during an expansion the pressure inside the cylinder only differs from that outside by an infinitely small amount, expansion is said to be ***reversible*** : it is important because the pressure p is then identical with the p occurring in the characteristic equation for the gas. Every isothermal expansion is reversible.

The equation to an isothermal is $pv =$ constant, for this is the type of expansion governed by Boyle's Law. It can be proved that the equation to a reversible adiabatic is $pv^\gamma =$ constant, where γ is equal to the ratio of the two principal specific heats of the gas, i.e. $\gamma = \frac{c_p}{c_v}$ [cf. p. 215].

Work Done by a Gas during a Reversible Expansion.—Suppose that a piston, acted upon by a pressure p from outside moves reversibly from A to B, Fig. 14·6 (a), between the walls of a cylinder containing gas. Such an expansion will take place when heat is supplied very slowly to the gas, and during the process the pressure of the gas inside the cylinder will remain constant: in fact it is equal to p. If the distance through which the piston moves is $(x_2 - x_1)$ and S is the area of the piston, the work done ***by the gas*** is equal to the force on the piston multiplied by the distance through which the point of application of the force moves; viz. $p\text{S}\,(x_2 - x_1) = p \times$ change in volume. If the change in volume is small, viz. δv, the work done by the gas in this small reversible expansion is $p.\delta v$.

Consider now the reversible expansion of a gas along a continuous curve AB, Fig. 14·6 (b) : let a and b be two points very close together

on the curve while m and n are the projections of these points on OX, just as M and N are the projections of A and B. The work done during the expansion from m to n is $am.mn$, since $abnm$ may be regarded as a rectangle when mn is small as we have assumed. It therefore follows that W the total work done by the gas during the reversible expansion from A to B is represented by the area ABNM, i.e. $\int_{v_1}^{v_2} pdv$.

This expression can be evaluated when the working substance is an ideal gas and the expansion is isothermal or adiabatic and

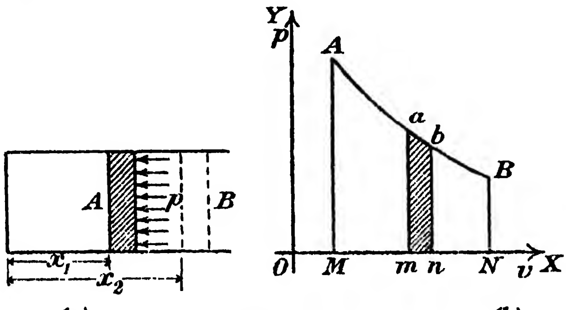


FIG. 14·6.

reversible. In the former instance, if we consider unit mass (1 gm.) of the substance

$$W = \int_{v_1}^{v_2} p\,dv = \frac{1}{\mathcal{R}T}\int_{v_1}^{v_2}\frac{dv}{v} = \frac{1}{\mathcal{R}T}\log_e\frac{v_2}{v_1}$$

$$= \frac{2{\cdot}303}{\mathcal{R}T}\log_{10}\frac{v_2}{v_1}.$$

For an adiabatic expansion from v_1 to v_2, we have $p_1v_1^\gamma = p_2v_2^\gamma$, so that

$$W = \int_{v_1}^{v_2} p\,dv = p_1v_1^\gamma\int_{v_1}^{v_2}\frac{dv}{v^\gamma}$$

$$= \frac{p_1v_1^\gamma}{\gamma - 1}\left[\frac{1}{v_1^{\gamma-1}} - \frac{1}{v_2^{\gamma-1}}\right]$$

$$= \frac{p_1v_1 - p_2v_2}{\gamma - 1} \qquad [\because\ p_1v_1^\gamma = p_2v_2^\gamma].$$

[For another expression for the work done in a reversible adiabatic expansion, cf. p. 288.]

On the Changes in Temperature Produced when a Gas suddenly Expands or Contracts.—When a gas is allowed to expand reversibly and no heat is permitted to enter the system there must

be a fall in temperature, for energy equal to the work done must be supplied by the molecules, i.e. their kinetic energy is reduced. Such an expansion is termed a ***reversible adiabatic*** or ***isentropic*** expansion. In the experiments to be described these conditions are only fulfilled approximately. A, Fig. 14·7 (*a*), is a wide glass tube fitted with a rubber bung through which passes a thermocouple connected in series with a high resistance galvanometer, G. A smaller tube leads from A to the large bottle B, in which the air pressure may be reduced with the aid of a filter pump as indicated. C is a stop-cock by means of which the air in A may be shut off from that in B, while D permits the whole apparatus to be filled with air at atmospheric pressure.

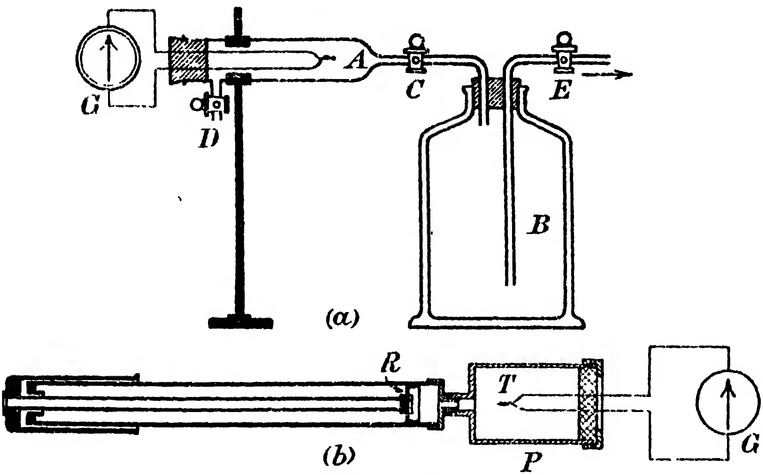


FIG. 14·7.—(*a*) Cooling Produced when a Gas suddenly Expands. (*b*) Heating Produced when a Gas is suddenly Compressed.

When the pressure of the air in B has been reduced and the tap E closed, C is opened—there is a sudden drop in the pressure of the gas in A and the galvanometer is deflected. After a short time this deflexion is reduced to zero showing that the temperature of the gas has been restored to its original value. To determine whether or not the gas was cooled or heated by the expansion, a lighted match is held below the thermocouple. The deflexion will be in the opposite direction to that cited above, i.e. the expansion of the gas was accompanied by a drop in temperature.

To show that there is a heating effect when a gas (air) is suddenly compressed, a metal tube, P, Fig. 14·7 (*b*), is attached to the end of a bicycle pump. A thermocouple, T, is placed inside this tube and connected to a high-resistance galvanometer, G. This thermo-element is supported in a rubber bung securely fastened to the

apparatus so that it shall not be forced out when the pressure inside is increased. The pressure of the air in the apparatus is suddenly increased by pushing in the piston R, and the deflexion of the galvanometer indicates that a heating effect has occurred.

[A mercury thermometer must not be used in these experiments owing to the change in volume of the bulb which takes place when the external pressure on it is varied.]

To account for this rise in temperature which occurs when a gas is suddenly compressed we have to use the principle of the conservation of energy, for the work done on the gas is equal to the increase in the kinetic energy of the molecules plus the energy which escapes from the system. Now under the conditions of the experiment this escape of energy is prevented so that the work done on the system is spent in increasing the energy of the gas molecules, i.e. the gas gets hotter. [The lower end of a bicycle pump becomes very warm when the pump is used to inflate a tyre.]

In a similar way it may be shown that a gas becomes cooler when it expands adiabatically and reversibly.

The Cooling Produced when an Ideal Gas expands Reversibly and Adiabatically.—Suppose that the initial state of unit mass of a gas is represented by the point A, (p_1, v_1), on a pv diagram—cf. Fig. 14·8. Let the gas expand reversibly and adiabatically to B, (p_2, v_2); the gas then receives heat from the surroundings until the pressure becomes p_3, the volume remaining v_2, while the temperature is restored to its initial value—its state is represented by C on the diagram.

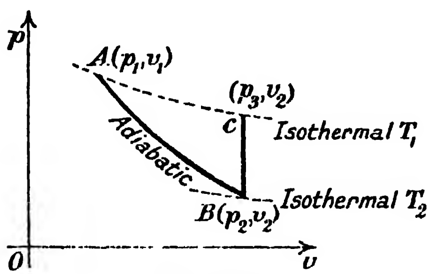


FIG. 14·8.

Since A and B are points on a reversible adiabatic

$$p_1 v_1^{\gamma} = p_2 v_2^{\gamma}.$$

But $p_1 v_1 = RT_1$ and $p_2 v_2 = RT_2$, since A and B are on the isothermals T_1 and T_2 respectively.

Eliminating the p's from these equations [we could, of course, eliminate the v's], we have

$$\left(\frac{v_2}{v_1}\right)^{\gamma} = \frac{p_1}{p_2} = \frac{T_1}{v_1} \cdot \frac{v_2}{T_2}.$$

$$\therefore \frac{T_1}{T_2} = \left(\frac{v_2}{v_1}\right)^{\gamma - 1}.$$

Since $v_2 > v_1$ and $(\gamma - 1)$ is positive it follows that $\left(\frac{v_2}{v_1}\right)^{\gamma-1} > 1$, so that when a gas expands reversibly and adiabatically cooling must occur. If T_1, v_1 and v_2, together with γ, are known, T_2 may be calculated and the cooling $(T_1 - T_2)$ deduced.

In consequence of the fact that the temperature falls when a gas expands reversibly and adiabatically it follows that the curve AB, Fig. 14·8, must be below AC, and if AB and AC represent the portions of the curves traced out in infinitely small expansions the slope of an adiabatic is necessarily greater than that of an isothermal at a point where the two curves intersect.

Calculation of J from the Difference between c_p and c_v.—Let *unit* mass of gas at a pressure p dyne. cm.$^{-2}$ and absolute temperature T occupy a volume v_1 cm.^{3}. To raise its temperature to $(T + 1)$ requires c_v cal. of heat if the volume remains constant. This heat is utilized entirely in increasing the kinetic energy of the gas molecules, if the gas is an ideal one. Now suppose that the unit mass of gas is heated from T to $(T + 1)$ at constant pressure. The heat required is c_p. But if the volume has changed from v_1 to v_2, some of the heat supplied will have been utilized in doing work against forces due to the external pressure. This work is $p(v_2 - v_1)$ [cf. p. 216]. Thus the heat supplied to increase the energy of the molecules in this instance is $c_p - \frac{p(v_2 - v_1)}{J}$. Now an equal increase in the molecular energy is obtained by supplying an amount of heat c_v at constant volume. Hence.

$$c_p - c_v = \frac{p(v_2 - v_1)}{J} = \frac{\mathscr{R}(T + 1) - \mathscr{R}T}{J} = \frac{\mathscr{R}}{J}.$$

This equation may be written

$$c_p - c_v = \frac{\mathscr{R}}{J} = \frac{1}{J}\left(\frac{pv}{T}\right),$$

so that $c_p - c_v$ may be calculated when p, v, and T are known. Again, we may write also

$$c_p - c_v = \frac{R}{JM},$$

where R is the gas-constant for one gram-molecule (the so-called universal gas constant) and M is the molecular weight of the gas considered.

Now for air at 0° C., $c_p = 0{\cdot}240$ cal. gm.$^{-1}$ deg.$^{-1}$ C.,

$c_v = 0{\cdot}171$ cal. gm.$^{-1}$ deg.$^{-1}$ C., and since air at S.T.P. has a density $0{\cdot}00129$ gm. cm.$^{-3}$

$$\mathcal{R} = \frac{p_0 v_0}{T_0} = \frac{76 \times 13{\cdot}59 \times 981}{273 \times 0{\cdot}00129} \text{ erg. gm.}^{-1} \text{ deg.}^{-1} \text{ K.}$$

$$\therefore J = 4{\cdot}17 \times 10^7 \text{ erg. cal.}^{-1}.$$

In the same way, if a gram-molecule of an ideal gas is considered, it may be shown that

$$C_p - C_v = \frac{\mathcal{R}}{J}$$

Since 1 gram-molecule occupies 22,400 cm.^{3} at S.T.P.

$$\mathcal{R} = \frac{p_0 V_0}{T_0} = 8{\cdot}34 \times 10^7 \text{ erg. gm.-mol.}^{-1} \text{ deg.}^{-1} \text{ K.}$$

$$= 2 \text{ cal. gm.-mol.}^{-1} \text{ deg.}^{-1} \text{ K.}$$

Thus the two principal molecular heats of an ideal gas differ by 2 cal. gm.-mol.$^{-1}$ deg.$^{-1}$ K.

In making the above calculation we have assumed that no work has been necessary to pull the molecules apart against their mutual attractions. To test this point JOULE (1845) devised the apparatus shown in Fig. 14·9 (*a*). A and B were two copper vessels connected together by a pipe fitted with a metal stop-cock of special design so that it was air-tight even when subjected to large pressure differences. A was filled with dry air at a pressure of 22 atmospheres, while B was exhausted. Both vessels were immersed in a large tank of water and when the temperature of the whole had reached a steady value, the stop-cock was opened. After thoroughly stirring the water no change in its temperature could be detected although the mercury thermometers used were sensitive to a change in temperature of 0·005° F. Joule concluded that no internal work was done. He then devised a modification of the above experiment and in doing so was actually repeating some earlier work by Gay-Lussac, although he was probably unaware of this fact. The above apparatus was inverted and the two copper vessels placed in different vessels of water, the thermal capacities of these vessels and their contents being known, the stop-cock also being immersed in another vessel containing water. When the gas was allowed to expand from A to B the temperature of the water round A fell appreciably, but this was accompanied by a rise in temperature of the water round B. The heat lost in A was very nearly equal to that generated in B. The very small inequality between these values vanished within the limits of accuracy of the experiment when a correction was made for the heat exchange between certain portions of the apparatus and their surroundings which we have neglected.

Now the cooling in the first vessel is due to the mechanical energy spent by the gas remaining in A in driving out that which has passed into B, and the heating in B is due to the work done on the gas already in that cylinder as this is compressed by the successive portions of gas which enter.

In the first experiment no mechanical work has been done on the whole and in its final state the gas occupies a volume twice as great as it did initially. It therefore follows that if there is any force of attraction between neighbouring molecules it is exceedingly small, i.e. the amount of internal work done by a gas during expansion is zero within the limits of accuracy of these experiments.

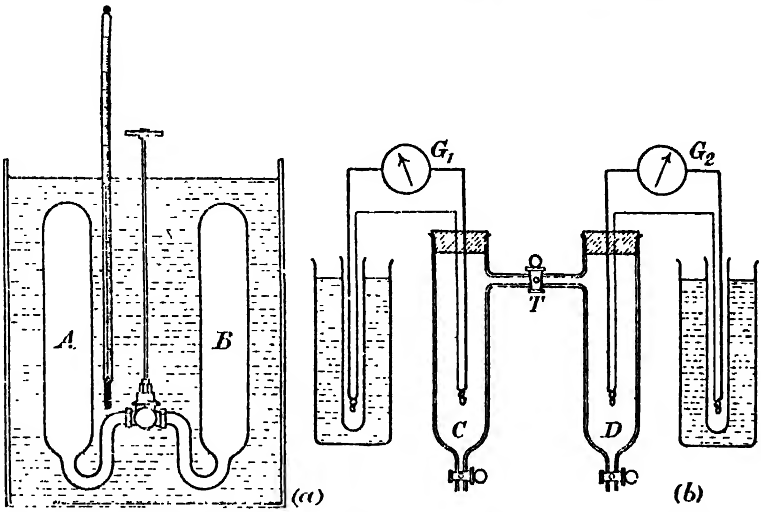


FIG. 14·9.—(a) Joule's first Apparatus for Investigating the Internal Work done by a Gas during Expansion. (b) Modern Form of Joule's Second Form of the above Apparatus.

Later work by JOULE and KELVIN showed that the internal work done when a gas expands is not zero, i.e. there is a definite force of attraction between the molecules of a gas and this is in part responsible for the deviations from Boyle's and Charles' Laws exhibited by all gases.

The heating and cooling effects obtained in the second experiment just described may be shown in the laboratory with the aid of the following apparatus :—It consists of two large thick-walled vessels, C and D, Fig. 14·9 (b). Each is fitted with a rubber bung through which one junction of a copper-constantan thermocouple passes. The other junctions belonging to each couple are placed in narrow glass tubes immersed in water at room temperature. G_1 and G_2 are high-resistance galvanometers placed in

series with each thermocouple and their deflexions serve to measure the difference in temperature between the junctions. D is exhausted and after several minutes, when the temperature is the same at all points in and near the apparatus, the stop-cock T is suddenly opened. The deflexions of G_1 and G_2 indicate that cooling and heating effects have occurred in C and D respectively.

An Ideal Gas.—This is defined as a gas which obeys the laws of Boyle and of Charles and for which there is no force of attraction between neighbouring molecules. This last part of the above statement signifies that the energy of the molecules is entirely kinetic and depends only upon the temperature of the gas.

Internal Energy and the Work Done in Reversible Adiabatic Expansions.—Consider one gram of an ideal gas. Let U be the energy associated with that mass of the gas when its temperature is T on the absolute scale. Then if δq is the heat added at constant volume so that no work is done, the energy imparted to the gas is $J\delta q$, and this must be equal to the increase in internal energy, i.e. δU.

$$\therefore\ J\frac{dq}{dT} = \frac{dU}{dT}.$$

But $\dfrac{dq}{dT} = c_v$, the specific heat at constant volume.

Hence $\delta U = Jc_v\delta T$.

When the volume of the gas does not remain constant the first law of thermodynamics gives

Heat (in ergs) added = increase in internal energy + the external work done by the gas.

Hence, in a reversible adiabatic expansion

O = increase in internal energy + external work done by the gas, and in an infinitesimally small such change this becomes

$O = Jc_v dT$ + external work done.

$\therefore$ In a finite reversible adiabatic expansion

$$O = J\int c_v dT + \text{external work done (W)}.$$

$$\therefore\ W = -J\int c_v dT = -Jc_v\int dT,$$

if c_v is constant.

$$\therefore\ W = -Jc_v(T_2 - T_1) = Jc_v(T_1 - T_2),$$

i.e. the work done per unit mass is equal to Jc_v times the fall in temperature. Since $\dfrac{c_p}{c_v} = \gamma$, and $c_p - c_v = J\mathscr{R}$, the above expression is identical with that obtained on p. 282.

Note on Experimental Methods for finding c_p.—In the experiments described on pp. 216 and 218, there is a flow of gas through the calorimeter: there must therefore be a difference in pressure between any two sections across the tube through which the gas flows. Do the methods therefore really give us values for c_p? The following argument is due to SEARLE. Consider unit mass of gas. Let p and v be the pressure and volume respectively, the suffix 1 denoting

conditions as the gas enters the flow tube, the suffix 2 conditions when it leaves. Then the work done per unit mass in forcing the gas into the entrance of the tube is p_1v_1; the work done against the atmosphere as the gas leaves the tube is p_2v_2. Let Q, in heat units, be the heat given to the calorimeter. Then

p_1v_1 + internal energy of gas on entering

= p_2v_2 + internal energy on leaving + heat energy given to calorimeter,

i.e.
$$p_1v_1 + U_1 = p_2v_2 + U_2 + JQ,$$
where J is the mechanical equivalent of heat. If the gas is an ideal one, $p_1v_1 = \mathscr{R}T_1$, and $p_2v_2 = \mathscr{R}T_2$, where T_1 and T_2 are the absolute temperatures of the gas entering and leaving the calorimeter, respectively. Moreover, $U_2 - U_1 = Jc_v(T_2 - T_1)$, where c_v is the specific heat of the gas at constant volume. Hence
$$\mathscr{R}(T_1 - T_2) = Jc_v(T_2 - T_1) + JQ.$$
But
$$\mathscr{R} = J(c_p - c_v).$$
$$\therefore J(c_p - c_v)(T_1 - T_2) = Jc_v(T_2 - T_1) + JQ,$$
$$\therefore c_p = \frac{Q}{T_1 - T_2} = \frac{Q}{t_1 - t_2},$$
where t_1 and t_2 are the temperatures recorded, i.e. c_p is actually measured although the gas is not at a constant pressure during the experiment.

The Isothermals for CO_2.—The relationship between the pressure and volume of a vapour is clearly indicated in the classical experiments made in 1869 by ANDREWS with carbon dioxide. The substance was contained in a thick-walled glass tube as indicated in Fig. 14·10 (*a*). The part AB was a fine capillary whilst BC was about 2·5 mm. and CD about 1 mm. in diameter. The tubes were calibrated by measuring the length of a thread of mercury at various positions in them and then determining the mass of the mercury used. Initially both ends of the tube were open and the gas was passed through for twenty-four hours in the hope that all traces of air would thereby be removed. The upper end of the tube was hermetically sealed and its lower end placed under mercury. Some gas was expelled from the tube by gently heating it, so that when it cooled a pellet of mercury was drawn into the tube. To adjust the quantity of gas in the tube to any desired amount the tube with its end D still below mercury was placed in a vessel connected to an air pump. When the pressure in the vessel was reduced some carbon dioxide escaped from the experimental tube. A similar tube was then filled with air. The two tubes were then enclosed in a copper vessel completely filled with water. Screw plungers in the base of the apparatus enabled the pressure on the gas to be increased. By observing the change in volume of the air Andrews deduced the pressure to which the gases were subjected. These calculations were made on the assumption that air was an ideal gas. Andrews showed that the tubes did not suffer a permanent enlargement even when subjected to high internal pressures

for some time. It appears from the original paper by Andrews that air separated the mercury pellet from the water so that the carbon dioxide never became moist. Similar conditions held in the other tube. No observations were obtainable at pressures less than 40 atmospheres since the mercury in the air-tube did not come into view until that pressure was attained. The air was kept at room-temperature: the capillary tube containing the carbon dioxide was kept at the desired temperature with the aid of a suitable jacket. Fig. 14·10 (*b*) shows the general nature of the curves exhibiting the relationship between pressure and volume for carbon dioxide at different temperatures.

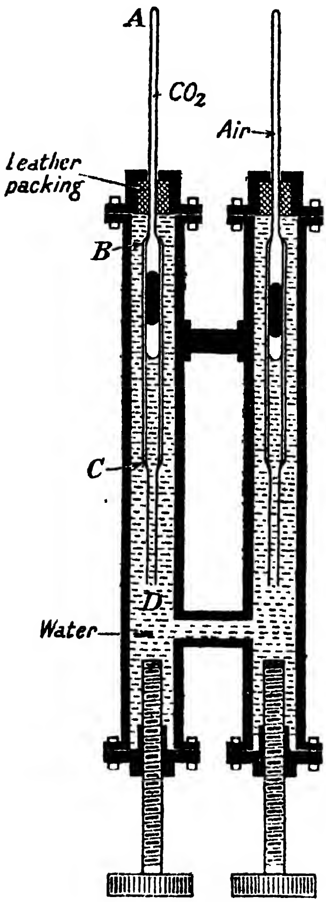


Fig. 14·10 (*a*).

It will be noticed that when the temperature is high the curves approximate to hyberbolæ, i.e. the gas behaves as an ideal gas approximately. At lower temperatures marked deviations become apparent. In the 31·1° C. isothermal a very short portion is horizontal and for all isothermals below this two sudden breaks appeared in each curve. The horizontal portion of the curve for 31·1° C. indicated that liquefaction had taken place. The particular temperature, 31·1° C., is termed the ***critical temperature*** for carbon dioxide, since at temperatures above this it is impossible to liquefy the gas. Strictly speaking, a substance should only be termed a ***gas*** when it exists at a temperature above its critical temperature. At lower temperatures it is a ***vapour,*** then a liquid, and finally a solid.

Any substance may therefore be a gas or a vapour; but a gas cannot be changed into a liquid by pressure alone, whereas a vapour may be changed into the liquid state at the same temperature by increasing the pressure, and during the transformation will exist as a saturated vapour in contact with its liquid.

The above facts are made more striking when the dotted curve shown in Fig. 14·10 (*b*) is drawn. Its left-hand branch is the locus of points for which a transformation from liquid to vapour begins, when a substance is heated at constant pressure; its right-hand branch is the locus of points at which vaporization is complete. These two branches are known as the ***liquid line*** and ***vapour line*** respectively, and, together, constitute the so-called ***border curve.*** These lines meet at C and touch the isothermal for the critical temperature at that point. Above C the isothermal for the critical temperature indicates the boundary between the two states. When the state of a substance is represented by any point within the

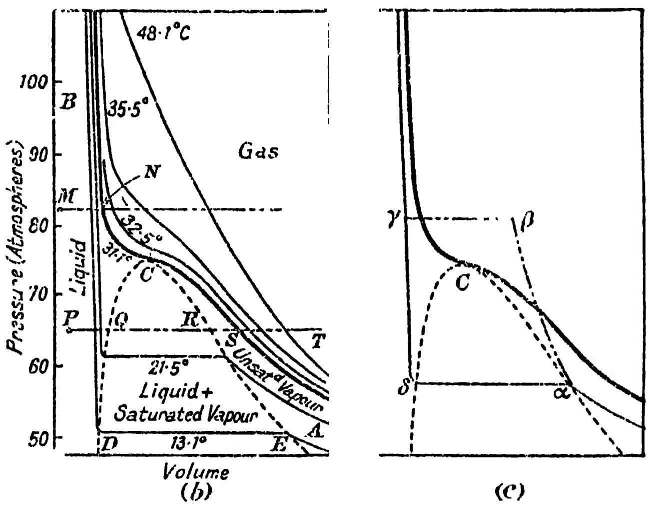


FIG. 14·10 (*b*) and (*c*).—Isotherms for CO_2.

border curve, then that substance exists partly as a saturated vapour and partly as a liquid.

The point C on the above diagram is known as the ***critical point;*** the pressure corresponding to this is the ***critical pressure,*** while the corresponding volume is the ***critical volume.*** Here we must remind ourselves that the curves we have drawn are for unit mass of the substance, so that the critical volume is the volume of unit mass of substance at its critical temperature and pressure. At the critical point, the densities of the liquid and vapour are equal.

At the point A on the diagram the carbon dioxide exists as an unsaturated vapour, whereas at B it is a liquid. By moving from A to a point opposite B in a direction parallel to the pressure axis, i.e. by heating the substance at constant volume, and then moving to B

along a line parallel to the volume axis, i.e. by cooling the substance at constant pressure, a transition from the gaseous to the liquid state will have been effected without any sudden discontinuity of state occurring.

[The fact that the curves are slightly rounded at those points where the whole of the carbon dioxide became liquid probably implies that all traces of air had not been removed from the carbon dioxide.]

Determination of the Critical Temperature and Critical Pressure of a Substance.—By heating a suitable quantity of the liquid under investigation in a sealed tube, the critical temperature of the substance may be ascertained by observing the tempera-

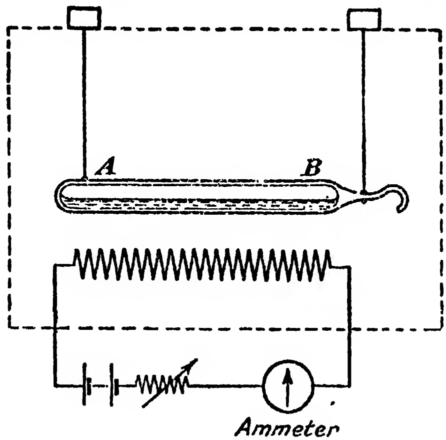


FIG. 14·11.—A Boyla Tube.

ture at which the surface of separation between the liquid and its vapour disappears, for when the temperature of a liquid is raised its surface tension decreases, the meniscus becomes more flat and disappears altogether when the critical point is reached. These phenomena, together with others, are observed when some ether, placed in a thick-walled tube—the so-called 'Boyla tube'—is heated by placing it above a wire carrying an electric current. This method of heating is desirable since the rate at which energy is supplied may be controlled easily. The tube AB, Fig. 14·11, is supported in a horizontal position between two vertical plates of glass which protect the experimenter if the tube should happen to explode. The space above the ether contains only the vapour of that liquid, so that the pressure inside the apparatus is equal to the saturation vapour pressure of ether at the temperature of the liquid. When the temperature is raised evaporation takes place

and the density of the vapour increases while that of the liquid decreases. Evaporation proceeds without ebullition as the temperature is raised until at a definite temperature a striking phenomenon is witnessed. The meniscus becomes ill-defined and finally disappears, this stage being accompanied by the formation of a peculiar mist which is far from being quiescent. The pressure inside the tube is then equal to the critical pressure for ether: the temperature is the critical temperature. At a temperature slightly above this the tube is filled with a homogeneous substance—a gas. When the tube is permitted to cool, a mist appears in the tube and, spreading from the centre where it is first formed, completely fills the tube. A further slight cooling causes the mist to vanish; the lower half of the tube is then filled with liquid, the upper containing only the saturated vapour of the liquid.

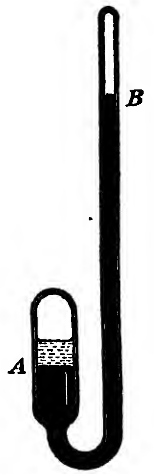


FIG. 14·12.—Caignard de la Tour's Apparatus for determining the Critical Pressure of a Liquid.

The above apparatus does not permit us to determine the critical pressure. CAIGNARD DE LA TOUR (1822-3) carried out the following classical experiment. The apparatus used consisted of a strong glass tube, AB, Fig. 14·12, filled with mercury from A to B, the space above A containing only the liquid and its vapour, while that above B contained air. From the volume of this air its pressure was calculated with the aid of Boyle's law, so that the pressure of the vapour above A became known. The liquid was heated and the phenomena described above were observed. The critical pressure for the substance under investigation was deduced from the volume of the air above B when the liquid in A just disappeared.

The Continuity of State.—In the ordinary process of evaporation or ebullition at constant temperature and pressure, the change of state from liquid to vapour occurs at the surface separating the two states. The process is a discontinuous one in the sense that all parts of the substance are not simultaneously affected in the same manner. The substance exists in two different forms, liquid and vapour in equilibrium with each other at the same temperature and pressure, and the whole process is characterized by the fact that the state of the substance is not homogeneous at any stage of the process. Let us examine whether or not a change of state may be brought about in such a way that the substance is homogeneous at all stages of the process.

Suppose that a liquid—say ether—is enclosed in a strong glass

tube in such a manner that it may be heated at constant pressure. Assume that pressure is less than the critical pressure. The isothermals will be similar in shape to those shown in Fig. 14·10 (*b*)—only the actual numbers will be different. Let P represent the initial state of the substance. Only liquid is present. When heat is supplied under the condition that the pressure remains constant, the changes which occur are represented on the above diagram by a straight line parallel to the volume-axis. Let this line intersect the border-curve at Q and R. Then at points on this line between Q and R both liquid and vapour are present—at R the substance will exist as vapour only. Suppose the heating is continued until a stage represented by the point S on the isothermal for the critical temperature is reached. During this stage the substance exists wholly as a vapour: when it is heated further its temperature is greater than the critical temperature and the substance is a gas.

If the constant pressure under which the heating takes place is greater than the critical pressure, suppose the initial state is represented by M. When the substance is heated under the stipulated conditions, it will exist as a liquid until the point N on the critical temperature isothermal is reached—it will then vapourize without any separation into two coexisting states occurring, i.e. there will have been no breach of homogeneity. In this way the transformation from liquid to gas, i.e. a vapour above the critical temperature for the particular substance investigated, may be effected by a continuous process without any breach of homogeneity. Theoretically, it is therefore possible to include both the liquid and vapour states in a single characteristic equation connecting the variables, p, v, and T; this equation represents an isothermal on a pv diagram for any given value of T.

An interesting series of transformations may be effected as follows: Suppose we begin with a saturated vapour corresponding to the point α Fig. 14·10 (*c*). By supplying heat and varying the pressure the point β may be reached. If the substance is then cooled at constant pressure to γ, it will have passed without ever existing in two different states simultaneously into a liquid. By releasing the pressure and abstracting heat, a point δ on the original isothermal may be reached. The substance is still liquid, but any addition of heat at constant pressure causes some vapour to appear.

The Liquefaction of Gases.—When the temperature or volume of an unsaturated vapour is reduced sufficiently the vapour becomes saturated, and if the reduction is continued some of the vapour will condense. About 1823 FARADAY succeeded in liquefying a number of gases. All the so-called gases are really unsaturated

vapours which may be liquefied by lowering the temperature and increasing the pressure to which they are subjected. To liquefy chlorine, for example, use is made of the fact that charcoal absorbs a large amount of this gas. A quantity of charcoal is saturated with chlorine and placed in one arm of a bent glass tube which is then closed at both ends and the other limb placed in a freezing mixture. When the charcoal is warmed gently, chlorine is evolved, and, when the pressure inside the apparatus is about 2 atmospheres, liquid chlorine appears in the cold limb.

While experimenting with carbon dioxide ANDREWS, as we have seen, discovered that it was impossible to liquefy this gas unless the temperature was below 31·1° C. however much the pressure was increased. It is found that all gases behave in this way, i.e. they cannot be liquefied unless they are ***first cooled below their critical temperatures.*** For a long time helium resisted all attempts to liquefy it and it was not until it was cooled to below −268° C. (its critical temperature) that liquid helium was obtained. By the rapid evaporation of liquid helium under reduced pressure a temperature of 1° K. has been obtained by ONNES. In 1934, F. SIMON, by a method depending on the demagnetization of a solid salt, reached a temperature of 0·1° K.

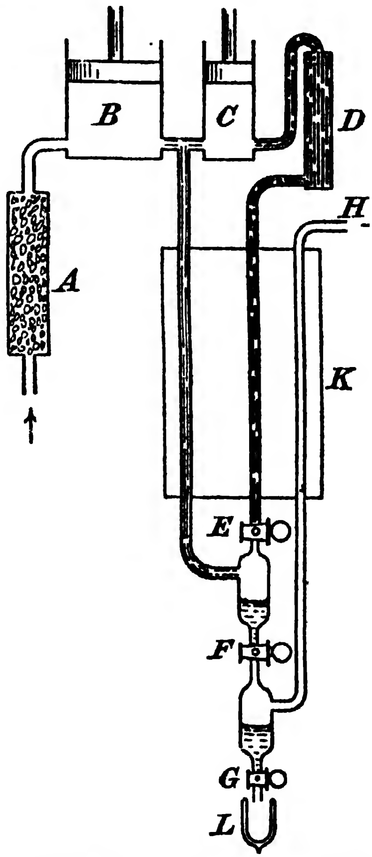


FIG. 14·13.—Linde's Apparatus for Liquefying Gases.

LINDE's apparatus for the liquefaction of air is shown diagrammatically in Fig. 14·13. On the upstroke of a piston, B, air is drawn over caustic soda contained in A so that the air is partly dry and free from carbon dioxide. [N.B.—The necessary valves are not shown.] On the down stroke of the piston this air is forced

into C where it is again compressed and passes through a tower D containing calcium chloride to remove the last traces of water. When the valve E is opened this air expands and cools itself. The cooled air passes back to C and the process is repeated. After a short time liquid air collects in the chamber below E and when the valve F is opened this may be run into the lower chamber. When the valve G is opened the liquid air passes into a Dewar flask L. Some of the liquid air in the lower chamber evaporates and escapes at H. This cold air passes through the jacket K and helps to cool the air under pressure.

EXAMPLES XIV

1.—An engine consumes 64 lb. of coal, the calorific value of which is such that when 1 lb. is burnt, 17·2 lb. of water at 100° C. can be converted into steam at the same temperature. The engine does 240×10^6 ft.-lb. of work. What percentage of the heat is wasted? [J = 1400 lb. deg. C. units when the heat required to raise the temperature of 1 lb. of water 1° C. is the unit.]

2.—A lead bullet at 15° C. strikes a target. If the lead is all just melted (325° C.), its specific heat being 0·031 cal. gm.$^{-1}$ deg.$^{-1}$ C., and its heat of fusion 5 cal. gm.$^{-1}$ determine the velocity with which the bullet hits the target. [$J = 4 \cdot 18 \times 10^7$ erg. cal.$^{-1}$.]

3.—Describe a method for determining the mechanical equivalent of heat. Obtain a value for the velocity with which a hailstone must strike the ground in order that, if three-quarters of its kinetic energy were converted into heat in the hailstone, one two-thousandth part of it would be melted.

4.—Calculate the difference in temperature between the water at the top and bottom of a waterfall assuming that 15 per cent. of the energy of fall is spent in heating the water which falls 25 metres.

5.—Draw a series of isothermal curves to show the relation between p and v at different temperatures for such a substance as CO_2. Point out from your diagram the distinction between a *gas* and a *vapour* and the meaning of *critical temperature*. Describe an apparatus by means of which the necessary data for drawing such curves may be obtained. (L. '31.)

6.—Explain two methods of producing low temperatures.

7.—Some ether is enclosed in a strong glass tube in such a manner that it can be heated at constant pressure. Describe and explain the phenomena which you would expect to occur when the ether is gradually heated to a temperature above its critical temperature, (*a*) if the pressure is less than the critical pressure, (*b*) if the pressure is greater than the critical pressure. Illustrate your answer by reference to a diagram of a series of isothermal curves for different temperatures.

CHAPTER XV

THE TRANSFERENCE OF HEAT—CONDUCTION AND CONVECTION

Conduction, Convection, Radiation.—Heat may be propagated from one point to another by three processes : (*a*) conduction, (*b*) convection, (*c*) radiation. In the processes of conduction and convection the molecules of the intervening matter are responsible for the transmission of the heat energy; according to the kinetic theory, the ultimate particles, or molecules, of a body are endowed with an ever-changing motion which becomes more vigorous as the temperature rises. In the process of conduction the molecules in the immediate vicinity of the source of heat become more vigorous, and impart energy by collision to their neighbours. These, in their turn, affect the next layer of molecules, and so the temperature tends to rise at all points in the conducting medium : it must be noted that the molecules of the body do not move along the body, but simply move about their mean positions. In the process of convection the heated molecules travel through the substance carrying their thermal energy with them, until it is lost by frequent encounters with more slowly moving molecules.

On the other hand, heat can be radiated through a vacuum (heat comes to us from the sun), a fact which at once proves that molecules of matter are not necessary for the transmission of heat by radiation. A heated body is a source of heat rays and these are propagated through space : when they impinge upon matter the molecules of that substance are excited, become more vigorous in their movement, and so there is an increase in temperature.

Moreover, while the process of heat transfer by conduction or convection is comparatively slow, radiant energy travels through space with the velocity of light; in bodies transparent to heat radiations, the velocity is somewhat less.

The Davy lamp, for use in mines, is a device for diminishing the risks of an explosion in the presence of combustible gases. Its success depends upon the fact that metals are good conductors of heat. Suppose a piece of metallic gauze is placed over a bunsen burner, the gas supply tap being turned on ; if the gas is lighted by means of a match placed above the gauze the flame is unable to penetrate below the latter. The reason for this is that before combustion of a gas can occur a certain minimum temperature is

necessary; the gauze is such a good conductor of heat that the temperature of the gas underneath the gauze is always below this minimum value and is therefore not ignited. In the Davy lamp a wire gauze separates the flame from the outside atmosphere, and so even in the presence of fire-damp the flame does not ignite the mine gases, because the temperature outside the lamp is below the critical one for combustion to take place. In the more modern forms of this lamp, Fig. 15·1, the wire gauze is arranged above the flame, the flame itself being surrounded by a glass cylinder to increase the illuminating power of the lamp. In addition, this more modern lamp may be used under conditions of greater danger, for, whereas the flame in the old lamp was forced through the meshes of the gauze when placed in an air current moving with a velocity of 5 ft. sec.$^{-1}$ the new pattern is safe even when the air velocity is 30 ft. sec.$^{-1}$

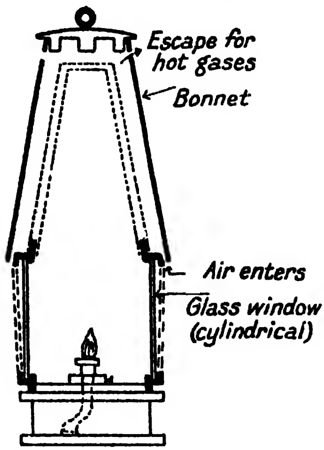


Fig. 15·1.—A Davy Safety Lamp.

The process of convection is utilized in the production of a hot-water supply for domestic and other uses—Fig. 15·2. The essential

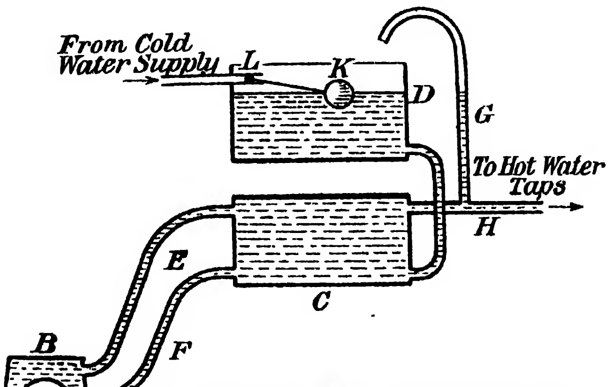


Fig. 15·2.—Domestic Hot-Water Supply.

parts are the hot-water cistern C and the boiler B, this latter being at a level below C. The hot water in B, having a lower density at higher temperatures, rises by way of the pipe E into C, its place being taken by cold water from C which enters B through the pipe F. As a result of this continuous process the water in the system

becomes heated. The tube H leads to the hot water taps, which must be below the tank D, which is a source of cold water. K is a float which falls when water is withdrawn from the system, thus permitting the valve L to open and water to enter through it. As the water enters so does K rise again until L is closed. The pipe G is a safety device which allows water to be forced through it, should it happen that the water boils.

The Sunvalve.—As an example of the use of radiant energy in operating a mechanical device let us consider the action of a sunvalve, which is used to control the supply of acetylene gas to beacon lights, etc., round the coast. It could properly be called a light valve, since it does not depend upon direct sunshine for its action, but only upon the degree of natural light prevailing. The principle underlying its operation is that a dull black or non-reflecting surface will become appreciably higher in temperature than a bright surface, when both are simultaneously exposed to light; though the temperature will become equal in the absence of all light. In the actual instrument the central cylinder is coated with lamp-black. Three gold-plated rods are arranged at intervals of 120° round this central column but at a little distance from it. The lower end of the central cylinder is provided with a needle-point bearing upon a pivoted horizontal steel tongue, which closes the gas outlet after the sunvalve has been exposed to daylight. After sunset the temperature of the valve becomes uniform, the centre cylinder contracts, the decrease in length being magnified by means of levers, and the gas outlet is opened. The great advantage of this instrument is that it is unaffected by climatic conditions, since applied heat or cold merely raises or lowers the temperature of the instrument as a whole, and does not occasion any relative displacement, as light does. A pilot light ensures that the combustion takes place when the valve is opened.

Thermal Conductivity.—It is a matter of everyday experience that some substances conduct heat more readily than others—silver is such a good conductor that the handle of a silver teapot must be insulated from the body by means of a poor conductor such as ebonite. Glass is such a badly conducting substance for heat that if hot water or milk is poured into a thick-walled or badly annealed glass vessel, the latter invariably cracks. This is because the heated portions of the glass expand and thereby produce strains in the glass which ultimately lead to fracture. Dewar vessels are double-walled glass vessels, the intervening space being a vacuum: a vacuum is the worst conductor of heat and, as a consequence, liquids placed inside the central vessel maintain their temperature for a considerable time—cf. p. 342.

Balsa wood, *Ochroma Lagopus*, is a very poor conductor of heat:

at the same time it is exceptionally light, its density being only 6 lb. ft.$^{-3}$, whereas that of mahogany is 45 lb. ft.$^{-3}$. This wood which is grown in Central America is utilized for the cold storage of fruit on ships. The poor conductivity of the medium helps in the maintenance of a low temperature, whilst its low density makes the freightage less. Another very poor conductor of heat is a form of rubber known as 'expanded' rubber. Its density is about 11 lb. ft.$^{-3}$ and it is an almost ideal badly conducting substance; its low conductivity arises from the fact that it consists of a very large number of minute cells all of which are filled with air.

Note on the 'Steady State.'—Suppose that a long bar, similar to that shown in Fig. 15·5 [cf. p. 304], has a number of small holes drilled in it so that the bulbs of mercury-in-glass thermometers may be inserted into them. Let one end of this bar, the whole of which is initially at room temperature, be heated. The temperature at any point near to the heated end soon begins to rise but a considerable time may elapse before the more distant parts of the bar receive an appreciable amount of heat. This is because as heat reaches any portion of the bar, say that between the two neighbouring planes normal to its length, some heat is utilized in raising the temperature of that section, some is lost from its sides, and the remainder passes on to the next section. Eventually a stage will be reached when the temperature of any such section remains constant—the so-called ***steady state*** will then have been reached. When this state exists at all points in the bar no heat is spent in raising the temperature of the bar, but that which reaches any section passes on to the next except in so far as some heat may be lost from the sides.

Thermal Conductivity.—To define the term 'thermal conductivity' let us imagine a small element bounded by two planes, PQ and RS, Fig. 15·3, each of area A, and normal to the direction

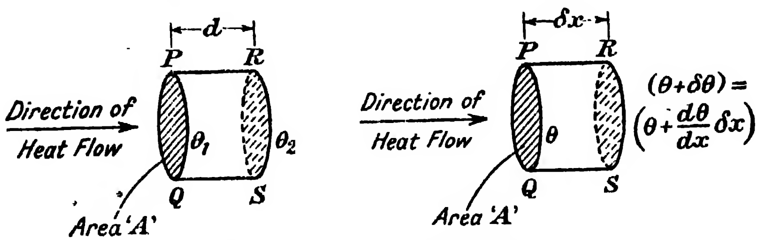


FIG. 15·3.—Definition of Thermal Conductivity.

of heat flow in a body where a 'steady state' exists. It is assumed that no heat escapes from the sides of the element. Let θ_1 and θ_2 be the temperatures of the faces PQ and RS respectively, $\theta_1 > \theta_2$.

If d is the distance between these planes, then $\frac{\theta_1 - \theta_2}{d}$ is the fall in temperature per unit distance in the direction of the flow of heat. If Q is the quantity of heat entering the element per second across the face PQ (and leaving it in the same time across RS) experiment shows that

$$\frac{Q}{A} \propto \frac{\theta_1 - \theta_2}{d}, \quad \text{or} \quad \frac{Q}{A} = \kappa . \frac{\theta_1 - \theta_2}{d},$$

where κ is a constant known as the thermal conductivity of the material.

[*Alternatively:* If θ and $\theta + \delta\theta$ are the temperatures at the faces PQ, RS, these being at distances x and $x + \delta x$ from a fixed plane in the material, then $\frac{d\theta}{dx}$ is the temperature gradient in the bar for the element considered, and under the conditions stated above

$$\frac{Q}{A} \propto - \frac{d\theta}{dx}, \text{ or } \frac{Q}{A} = - \kappa \frac{d\theta}{dx},$$

where κ is the thermal conductivity of the material.]

If H is the quantity of heat passing in t seconds, we have

$$\frac{H}{A} = \kappa . \left(\frac{\theta_1 - \theta_2}{d}\right)t.$$

κ is usually expressed in cal. $cm.^{-1}$ $sec.^{-1}$ $deg.^{-1}$ C.

If, now, we imagine PQ and RS, Fig. 15·3 (a), to be cross-sections of a well-lagged bar which has reached the steady state, Q will be constant for each normal section of the bar, so that the fall in temperature per unit distance along it is constant if κ is a constant. Experimentally, therefore, when such conditions have been established, it suffices to measure the temperatures at any two sections at a known distance apart in order to find κ when Q and A are known. If, however, heat escapes from the sides of the bar, the fall in temperature per unit distance along it is not constant, and the determination of κ becomes more difficult.

Searle's Apparatus for Determining the Thermal Conductivity of Copper.—This apparatus, shown in Fig. 15·4, may be used to determine the thermal conductivity of very good conductors of heat, such as silver and copper. In the latter instance it consists of a bar of copper 5 cm. in diameter and 40 cm. long, fitted at one end with a steam chamber A and at the other with a cooling chamber C. A copious supply of steam is passed into A through the wide tube B, and the water formed from the condensed steam passes back into the boiler without interrupting the flow of steam. A steady stream of cold water is passed through C via the inlet D and the exit E, and this water is caused to cir-

culate round the bar by means of a copper spiral whose coils have a thickness equal to the distance between the bar and the outside walls of C; the coil is extended so that the water passes freely. The apparatus is enclosed in a wooden box lined with felt. When steam is circulated through A the water flowing through C becomes heated, so that the heat passing along the bar in a given time is determined by finding the mass of water flowing in the same time and observing the rise in temperature of the water by means of the thermometers T_3 and T_4. This quantity will be equal to the heat entering the bar from the steam chamber in the same time

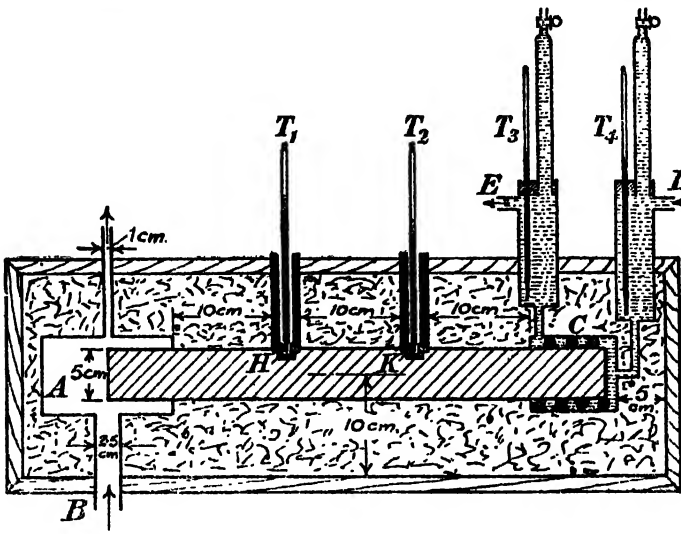


FIG. 15·4.—Searle's Apparatus [with modifications] for Determining the Thermal Conductivity of Copper.

providing that no heat escapes from the sides of the bar. The amount of heat lost from bars of good thermal conductivity is small and in the present instance it is further reduced by surrounding the bar with a badly-conducting substance, so that when steady conditions have been attained the amount of heat escaping from the sides of the bar is negligibly small.

To measure the temperature gradient along the bar, two thermometers, T_1 and T_2, are inserted into small holes drilled in the bar at a distance apart of 10 cm. These holes should not extend to the centre of the bar, for if they do the uniform flow of heat in the bar will be disturbed just at those points where the absence of any such disturbance is essential. To reduce this violation of uniform flow of heat along the bar, it is provided with small copper pieces H and K which are soldered to the bar and which

extend beyond the sides of the bar so that the bulbs of the thermometers are completely surrounded by metal and also so that narrow ebonite tubes may be fixed to them. These permit the thermometers to be inserted into their respective positions. In this way the temperatures of the sections through H and K are found without disturbing the heat flow seriously. If the readings of the thermometers T are denoted by θ with appropriate suffixes, and d is the distance apart of T_1 and T_2, the temperature gradient in the bar is $(\theta_1 - \theta_2) \div d$, providing that the surfaces of uniform temperature in the bar are at right angles to its axis. It is to attain this condition that in the apparatus here shown, the distances between T_1 and A, and between T_2 and C, have been doubled in comparison with the usual dimensions found in this apparatus, so that any want of uniformity in the temperature distribution near A will not affect the flow of heat in the region where the temperatures are observed.

If m is the mass of water flowing per second, the heat passing along the bar in this time is $m(\theta_3 - \theta_4) = Q$. But

$$\frac{Q}{A} = \kappa \text{ (fall in temperature per unit distance along the bar)},$$

where κ is the thermal conductivity of the material of the bar, and A the cross-sectional area. Hence

$$\frac{m(\theta_3 - \theta_4)}{A} = \kappa\frac{(\theta_1 - \theta_2)}{d}$$

so that κ may be determined.

The above theory further assumes that the indications of the thermometers are correct. To allow for any serious departure from this, T_1 and T_2 are placed in a bath at temperature θ_2 and with approximately the same lengths of column exposed as in the actual experiment. If T_1 then reads ϕ_2, the corrected temperature gradient is $(\theta_1 - \phi_2)/d$; similarly, if T_3 reads ϕ_4 when T_3 and T_4 are in a bath at temperature θ_4, the corrected rise in temperature of the water is $(\theta_3 - \phi_4)$.

The Thermal Conductivity of the Material of a Bar—Forbes' Method.—The material whose thermal conductivity is required is in the form of a bar about one metre long. The bar is curved at one end and this portion dips into a crucible containing molten lead or solder, Fig. 15·5. When this has melted, the gas flame is lowered and the temperature maintained at the melting-point of the metal by adding small pieces of the latter. Thermometers are inserted into holes drilled at various distances along the bar. These holes are filled with mercury (or fusible alloy—Wood's metal—at the hotter end of the bar on account of the high vapour-pressure of mercury at such temperatures) to improve the thermal contact between the bar and the thermometers. This

portion of the bar is protected from radiation from the hot bath by a suitable screen. Let us fix our attention on that portion of the bar between the points A and B. Heat is conducted through the hotter end A to this portion AB : part of this heat will be lost by conduction through the end B, part will be emitted from the surface of the bar, and in the initial stages of the experiment the remaining fraction of the heat entering the bar will be utilized in raising the temperature of AB. Eventually, however, a ***steady state*** is reached when the temperature no longer increases. When this happens the whole of the heat passing any cross-section of the bar in a given time is equal to the heat lost in the same time from the whole of the surface lying beyond the particular section considered.

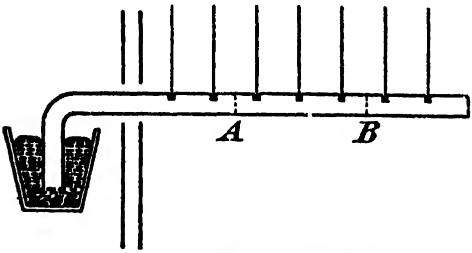


FIG. 15·5.—Forbes' Bar Apparatus.

When this state has been reached the temperatures of the thermometers are recorded and a curve showing the distribution of temperature along the bar is constructed (Fig. 15·6). If the tangent CD at a point C on this curve is drawn, the slope of this line gives the temperature gradient at a point in the bar corresponding to C. The thermal conductivity κ may then be calculated if the quantity of heat passing across the section at C per second can be estimated.

To do this a short piece of material having the same cross-section as the bar is heated to a temperature somewhat in excess of that of any portion of the bar on the side of the screen remote from the bath. This is then allowed to cool and a cooling curve constructed. The slope of this curve is a measure of the rate at which heat is lost from any portion of the bar when at the temperature corresponding to the point where the slope is measured. To calculate the heat lost by the bar in the static experiment the bar is imagined to be divided into short elements of the same length as the bar used in the second or dynamic experiment. The mean temperature of each such element having been ascertained from the curve in Fig. 15·6 the loss of heat from each is computed. The total heat radiated from the bar beyond C is therefore known, for it is equal to the sum of the heats lost by each element beyond that point.

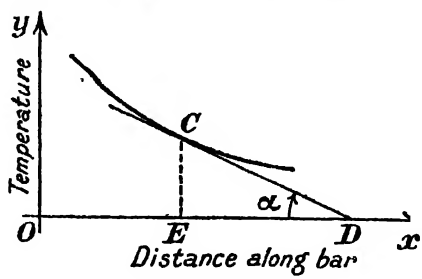


FIG. 15·6.

Distribution of Temperature along Lagged and Unlagged Bars heated at One End and in the Steady State.—We shall assume that the bars have constant cross-sectional areas and that

the thermal conductivities of their materials are constant. Consider the equation

$$\frac{Q}{A} = -\kappa\frac{d\theta}{dx}, \text{ where } \frac{d\theta}{dx} \text{ is the temperature gradient in the bar.}$$

If the bar is lagged, so that no heat escapes from the sides of the bar, it follows that the temperature gradient in the bar is constant and negative. The temperature distribution along the bar is therefore a linear one, the temperature falling as one recedes from the heated end of the bar—cf. AB, Fig. 15·7.

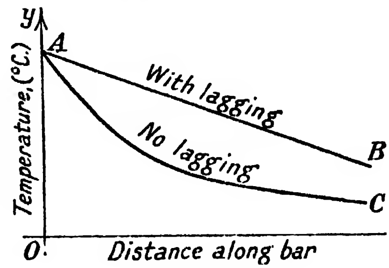


FIG. 15·7.

If, however, the bar is not lagged, but the steady state has been reached, a portion of the heat passing across any cross-section of the bar is lost from the sides, so that as we proceed along the bar away from the heated end smaller and smaller amounts of heat traverse consecutive sections. It follows that the numerical magnitude of $\frac{d\theta}{dx}$ becomes less and less—the temperature distribution being as in AC, Fig. 15·7.

[Before the steady state is reached the distribution is never a linear one, the temperature at any point always being less than the temperature at that point when the steady state has been reached, whether or not the bar is lagged.]

The Comparison of Thermal Conductivities.—INGEN HAUSZ is responsible for the following approximate method of comparing the thermal conductivities of two metals—say copper and bismuth. A, Fig. 15·8, is a metal tank in which water is kept boiling by an electric heater. B and C are rods of bismuth and copper, respectively, each 20 cm. long and 1 cm. in diameter. Both rods are electroplated and highly polished so that they shall lose heat at the same rate under the same conditions. Small lead shot are attached by means of paraffin wax at a regular distance apart to the under side of each bar. To obtain as much information as possible from this experiment it is advisable to arrange everything in position and then

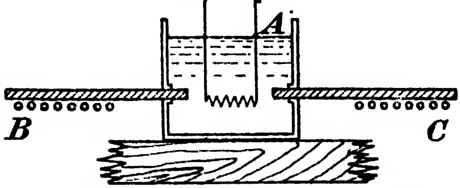


FIG. 15·8.—Ingen Hausz's Apparatus.

pour boiling water into A, which is kept boiling by the energy dissipated in the heating coil. As heat is conducted along the rods the wax melts and some of the shot fall off. It will be found that the shot become detached from the bismuth first, but this does not prove that the thermal conductivity of bismuth is greater than that of copper. The reason for this is that during the initial stages of the heat-flow along the rods the rate of rise in temperature depends not only on the thermal conductivity of the specimen but also on its thermal capacity, so that if the thermal capacity is small the initial rise in temperature of a poor conductor may be greater than that of a good conductor having a high thermal capacity.

The steady state eventually reached in this experiment occurs when the heat flowing across any section of the bar is equal to the amount emitted from the surface of the bar beyond that section. If the emissivities of the surfaces of the two bars are identical it can be shown that if κ_1 and κ_2 are the thermal conductivities of the two metals and l_1 and l_2 the distances from the hot end to the point where the wax just melts, then

$$\frac{\kappa_1}{\kappa_2}=\frac{l_1^2}{l_2^2}.$$

As arranged above, l_1 and l_2 are proportional to the number of shot which fall from each bar respectively.

The Thermal Conductivity of Mercury—Berget's Method.—The guard-ring method was applied by BERGET with considerable success to determine the thermal conductivity of mercury. His apparatus is shown in Fig. 15·9. AB is a glass tube surrounded by a wider tube, CD. Each is filled with mercury to the same level as indicated, the mercury in the outer tube serving as a guard-ring, i.e. this mercury prevents the loss of heat by lateral radiation so that the mercury in AB may be regarded as part of an infinite wall of mercury with its upper and lower faces at constant temperatures. To measure the heat passing down the column AB a Bunsen ice calorimeter was used. The column AB protruded into the central part of this instrument. The mercury guard-ring rested on an iron plate, P, while the calorimeter was surrounded by melting ice. The mercury was heated by passing steam through the tubes shown at the top of the diagram. In the final experiments made by Berget, the tubes through which the steam entered were almost in contact with the mercury surface and the supply was sufficiently rapid to blow to one side the water formed from the condensed steam. In this way the temperature of the upper surface was maintained at that of the steam.

If the thermal conductivity of mercury is independent of the

temperature, there will be a linear distribution of temperature along AB. This was investigated by means of thermocouples arranged as shown. Through small holes in the glass tubes iron wires covered with rubber were introduced, only the extremities of the wires being bare and in contact with the mercury. Any two of these wires and the mercury between them constituted a thermocouple. Berget found that the distribution of temperature along AB was linear and from the known dimensions of the apparatus calculated the heat conductivity of mercury. He obtained a value 0·0202 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C.

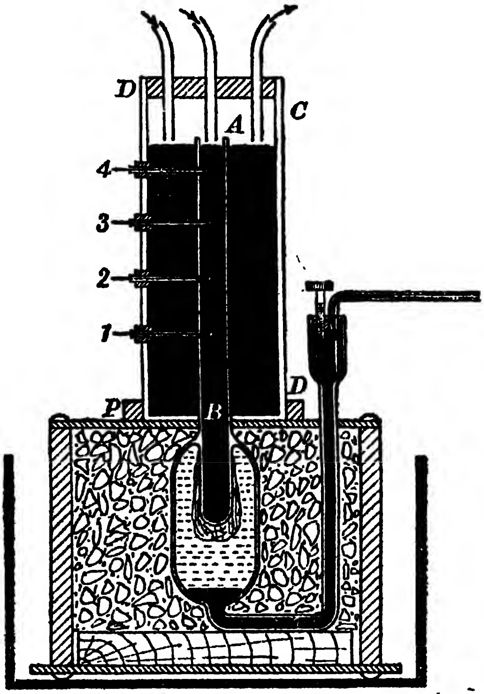


Fig. 15·9.—Berget's Apparatus for Determining the Thermal Conductivity of Mercury.

The results obtainable with this apparatus cannot be considered very accurate, since the conditions at the lower end of the guard-ring are not identical with those at the lower end of the column AB, for here the temperature cannot be 0° C., and an essential feature of a satisfactory guard-ring is that the conditions at its ends near to the ends of the column which is 'guarded' shall be identical with those at the ends of the above column.

It is interesting to note that Berget calculated what was the temperature of the upper surface of the mercury from the temperature gradient measured in the mercury column. In two experiments he found this to be 99·8° C. and 100·0° C. when the steam temperature deduced from the reading of the barometer was 100·1° C. and 100·4° C. The method of heating the mercury must therefore be considered satisfactory.

Callendar's Method for Rock Specimens and other badly conducting Substances.—The apparatus is shown in Fig. 15·10

and consists essentially of a heating coil B placed inside a gun-metal box about 5 in. square in section. D is the rock specimen, also 5 in. square in section and 2 in. thick, resting on top of the gun-metal box. On top of D is another square metal chamber, E, through which a rapid stream of cold water is passed. T_1 and T_2 are two thermocouples serving to measure the temperatures θ_1 and θ_2 at the faces of the specimen. These are insulated from B and E by very thin pieces of mica ($<$ 0·001 in.). A little paraffin wax is placed on each face of the rock and the apparatus raised to such a temperature that the wax melts when it is placed between clamps and pressed together so that, when

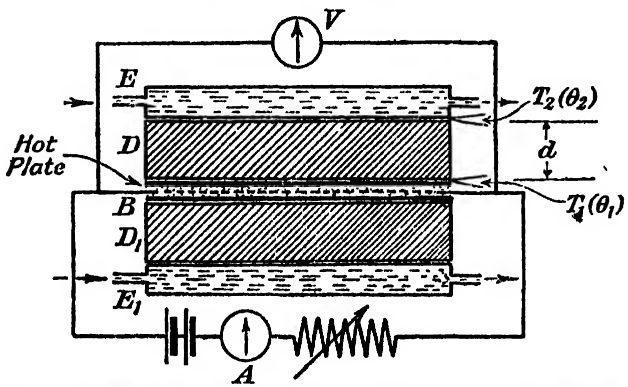


FIG. 15·10.—Callendar's Apparatus for Determining κ for Rock Specimens.

cold, the wax, although only about 0·001 in. thick, serves to make the apparatus rigid. The portions D_1 and E_1 of the apparatus are identical with D and E. The energy dissipated *per second* in the heater is VI joules, where V is the potential difference in volts across the coil and I the current through it in amperes. Since the apparatus is symmetrical about B, one half of this heat passes through each specimen when conditions have become steady, if we neglect the small quantity of heat lost by radiation, etc. The thermal conductivity κ is determined from the equation

$$\frac{VI}{2JA} = \kappa \cdot \frac{\theta_1 - \theta_2}{d},$$

where A is the cross-sectional area of the block and J the mechanical equivalent of heat in joules cal^{-1}.

Since the diameters of the wires constituting the thermocouples were less than 0·001 in. d was taken as the thickness of the specimen. It is necessary to have a rapid stream of water through E and E_1, and, in consequence, a very small rise in its temperature, so that

the temperatures of the outer faces of the specimen shall be uniform. Steady conditions are attained in two hours.

Lees' Disc Apparatus.—The thermal conductivity of a badly-conducting substance available in the form of a disc about 2 mm. thick may be determined by a method due to LEES. A simple form of this apparatus is shown in Fig. 15·11. A cylindrical slab of polished brass, A, is suspended in a horizontal position from a large metal ring, supported by a retort stand. Upon this rests a hollow cylinder, B, of the same diameter and provided with inlet and outlet tubes X and Y respectively. The base of this cylinder is similar to A. Mercury thermometers, T_1 and T_2, are inserted in holes bored radially in the base of B and in A respectively. A thin slab of the material under investigation—say ebonite—is inserted between A and B so that the space between the metal plates is completely filled with it. The metal part of the apparatus is nickel-plated to secure a uniform surface emissivity.

The hollow cylinder is first raised so that it is not close to A and a copious supply of steam passed through B. While the steam is still passing B is lowered on to the disc and the readings of the thermometers noted at half-minute intervals. Heat is conducted across the ebonite to A at a rate which slowly diminishes since the temperature difference between A and B decreases. When the temperature recorded by T_2 is about 70° C., the hollow cylinder is raised, the ebonite removed, and a cooling curve for the lower disc A obtained in the usual way. If we assume that all the heat passing into and through the ebonite disc is utilized in raising the temperature of the metal cylinder below it or radiated from the exposed surface of the cylinder, we may calculate the thermal conductivity of the ebonite as follows:

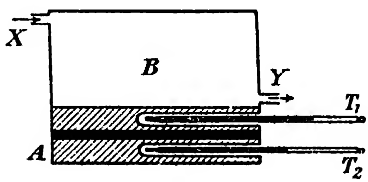


FIG. 15·11. **Lees' Disc Apparatus.** [Simplified Form].

Plot the heating and cooling curves for the lower cylinder and draw tangents to the curves at points having equal temperature co-ordinates. If α and β are the slopes the tangents to the heating and cooling curves make respectively with the time axis, we have,

$$\frac{\left[ms\,\alpha + ms\,\beta\left(\dfrac{\pi r^2 + 2\pi rz}{2\pi r^2 + 2\pi rz}\right)\right]}{A} = \kappa \cdot \frac{(\theta_1 - \theta_2)}{d},$$

where κ = thermal conductivity of ebonite,
A = cross-sectional area of ebonite disc,
d = thickness of ebonite,
r = radius of metal cylinder,
m = mass of metal cylinder,
z = thickness of metal cylinder,
s = specific heat of metal,
θ_1 = temperature of upper surface of ebonite at instant considered,
θ_2 = temperature of the lower surface of ebonite and is that temperature at which the tangents have been drawn.

The expression in square brackets is the heat passing per second through the ebonite, while the expression in round brackets is the ratio of the exposed surface of the metal cylinder in the actual experiment to that in the cooling experiment. Although this correction is applied it can only be approximate since in the cooling experiment more heat will be lost from the upper surface than from the lower.

Another uncertainty in the above expression arises from the fact that we have tacitly assumed that no heat is utilized in raising the temperature of the ebonite or radiated from its surface; but since the ebonite is thin and has a small thermal capacity compared with that of the metal cylinder this correction is small.

The above method of carrying out this experiment enables several values of κ to be determined from one set of observations; the following method, in which it is true that steady conditions are reached, so that no heat is spent in raising the temperature of the disc or cylinder but all is lost from the surfaces, possesses the disadvantage that only one estimate for κ can be made.

The heating is continued until steady conditions have been obtained. Then if ϕ_1 and ϕ_2 are the temperatures of the faces of the ebonite of thickness d, we have

$$\frac{Q}{A} = \frac{\kappa(\phi_1 - \phi_2)}{d},$$

where Q is the heat passing per second through the disc and also radiated from the surface of the cylinder in the same time. This is $ms\delta$, where δ is the rate of cooling when steady conditions have been obtained. To determine δ the ebonite is removed, and the heating continued so that the temperature of the lower cylinder is raised above ϕ_2; the heater is then removed and a cooling curve constructed. Then, if γ is the rate of cooling at ϕ_2, we have

$$\delta = \gamma\left[\frac{\pi r^2 + 2\pi rz}{2\pi r^2 + 2\pi rz}\right] = \gamma\left[\frac{r + 2z}{2r + 2z}\right]$$

as before, so that κ may be found. In this method any uncertainty in the relation between δ and γ affects Q to the same extent, so that the first method is preferable. Hence

$$\kappa. = \frac{ms\gamma(r + 2z)}{\pi r^2(2r + 2z)} \cdot \frac{d}{(\phi_1 - \phi_2)}.$$

Thermal Conductivity of Glass or Porcelain.—When the material under examination is a badly conducting substance, obtainable in the form of a tube, the following method is useful. The tube AB, Fig. 15·12, is surrounded by a wider tube through which steam is passed, the steam entering at E and escaping at D. A steady stream of water (obtained in the manner indicated) enters the tube at F and finds an exit at G, the initial and final temperatures being recorded by the thermometers T_1 and T_2. The T-pieces which enable these thermometers to be inserted in the water easily are jacketed with cotton wool so that the heat-content of the water at these two points shall be invariable. A narrow copper rod C is wound with a piece of rubber in the form of a very open spiral so that the temperature at any particular cross-section of the flow

shall be uniform. The flow of water is adjusted so that the difference in temperature between T_1 and T_2 is about 10° C., otherwise air bubbles will be expelled from the water and vitiate the conditions for steady flow: even if recently boiled distilled water is being used and no air bubbles are formed it is inadvisable for the rise in temperature to be much greater owing to the large heat losses which would accrue.

Steady conditions having been established, the temperatures are recorded and the mass of water flowing per second deduced by

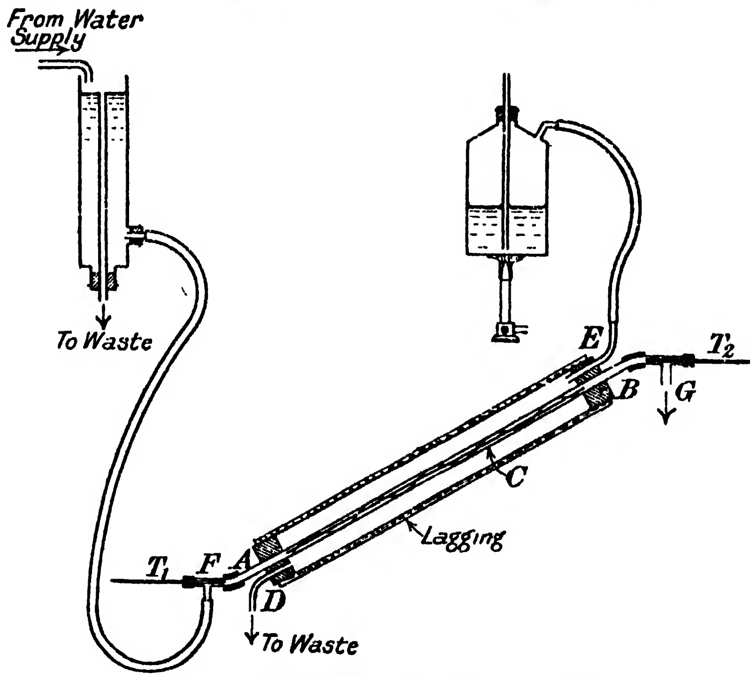


FIG. 15·12.—Thermal Conductivity of Glass.

observing the time in which a definite quantity of water is collected in a weighed conical flask. Let M be the mass of water flowing per second, θ_1 and θ_2 the initial and final temperatures of this water. Then M $(\theta_2 - \theta_1)$ is the quantity of heat passing per second through every co-axial cylindrical element of the tube. If l is the length of the tube, taken from the centre of one cork to that of the other, since we are uncertain regarding the effective length of the tube, r_1 and r_2 its internal and external radii respectively, ϕ_2 the temperature of the steam, and $\phi_1 = \frac{1}{2}[\theta_1 + \theta_2]$ the mean temperature of the inside wall of the tube, then $(r_2 - r_1)$ is the thickness and $2\pi \times \frac{1}{2}(r_1 + r_2) \times l$ the mean area of the material through which

heat is flowing. The thermal conductivity is therefore given by the equation

$$\frac{M(\theta_2 - \theta_1)}{\pi(r_1 + r_2)l} = \frac{\kappa[\phi_2 - \frac{1}{2}(\theta_1 + \theta_2)]}{r_2 - r_1}.$$

Before the steam is passed the readings of the thermometers are recorded—in general they will not be equal because no two mercury thermometers (at least the cheap ones found in laboratories) are consistent—and the correction to be applied to one of them in order to make its indication agree with that of the other is deduced. Steam is now passed, the readings of the two thermometers again being observed—the correction is applied to one of them and the true difference calculated—thus :—

Initial readings of the two thermometers X and Y are respectively 18·1° C. and 20·2° C.

$\therefore$ the correction to be applied to Y is $-$ 2·1° C.

Final readings of the two thermometers are 18·5° C. and 26·1° C. respectively.

$\therefore$ True final reading of Y is 24·0° C., so that the rise in temperature is 24·0 $-$ 18·5 $=$ 5·5° C.

A Guard-ring Method of determining the Thermal Conductivity of a Badly Conducting Substance.—The hot plate in this apparatus consists of two large copper plates with a resistance coil sandwiched between them. This is called the central hot plate. The coil is insulated from the plates by mica or micanite.

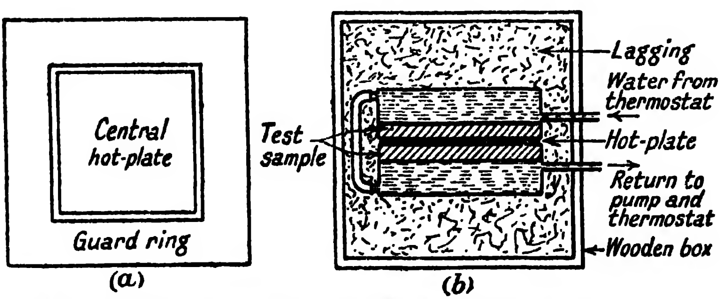


FIG. 15·13.—Guard-ring Method for investigating the Thermal Conductivity of a badly-conducting Substance (N.P.L.).

The guard-ring consists of an outer plate similar in construction to the above, but provided with an aperture into which the central hot plate may be inserted. There is a small clearance between the two plates. The surfaces of the plates are coplanar. The function of the guard-ring is to eliminate edge effects and ensure that the flow of heat from the central hot plate is normal

to its surfaces—cf. Fig. 15·13 (*a*). Two slabs of the material to be investigated are required and these are placed above and below the heating element—cf. Fig. 15·13 (*b*). The slabs must be equal in cross sectional area to that of the guard-ring. The apparatus is provided at the top and bottom with chambers through which rapid streams of water at a constant temperature flow.

Electrical energy is then dissipated in the central hot plate and in the guard-ring, each being controlled separately. Copper-constantan thermocouples are made by soldering a copper and a constantan wire to a very thin and small piece of copper: the copper discs are distributed over the surfaces of the heating plates and the dissipation of energy adjusted until the temperature is the same over the inner portion of the ring as it is over the central plate. A week may elapse before steady conditions are obtained.

The temperature difference across each slab is then measured by other thermocouples, and the thermal conductivity calculated as follows:

Let W be the rate at which energy is dissipated in the central hot plate—this is in watts and is equal to VI, if V is the voltage across the central heating coil and I the current in amperes through it. Let A be the cross-sectional area of the central plate—this is also the area across which the energy dissipated in the central plate flows. Let d_1 and d_2 be the thicknesses of the slabs, θ_1 and θ_2 the temperature differences across them. Then

$$\frac{Q}{A} = \frac{W}{JA} = \frac{VI}{JA} = \kappa\left[\frac{\theta_1}{d_1} + \frac{\theta_2}{d_2}\right],$$

where J is the mechanical equivalent of heat, expressed in joule. cal.$^{-1}$

The above is a short account of a precision method developed at the National Physical Laboratory, Teddington, by Ezer Griffiths.

The Flow of Heat across Composite Plates.—Let us consider the heat flowing across a portion of a large wall of a room consisting of a thickness of brick covered with plaster. Let the thickness of the plaster be d_1, while d_2 is that of the layer of brick: let κ_1 and κ_2 be the mean thermal conductivities of the plaster and brick respectively. Let θ_1 and θ_2 be the temperatures inside and outside the room ($\theta_1 > \theta_2$). Let A be the area across which the heat flow is considered. This area must be such that the heat flow is normal to the surfaces of the materials so that our equation may be applied. If Q is the quantity of heat flowing per second, when the steady state has been reached, across any section of the portion of the wall chosen parallel to the faces—then Q is constant for

all such sections, since there is no accumulation of heat at any point —we have, if θ is the temperature at the brick-plaster interface,

$$\frac{Q}{A} = \kappa_1 \frac{\theta_1 - \theta}{d_1} = \kappa_2 \frac{\theta - \theta_2}{d_2}.$$

If κ_1 and κ_2 are known, θ, and then Q, may be calculated.

The Thermal Conductivity of Water.—Liquids are poor conductors and measurements of their conductivities are rendered difficult on account of the presence of convection currents. That water is a poor conductor is shown by the fact that it may be boiled at the top of a large test-tube even while a piece of ice remains at the bottom of the tube—the ice must be weighted, say with wire gauze, to make it sink.

The Thermal Conductivity of Liquids.—The following apparatus has been designed by the author for determining the thermal conductivity of a liquid. The essential parts of the apparatus are shown in Fig. 15·14 (*a*). It consists of a square hot plate, A, made by sandwiching a heating mat between two brass plates. The mat consists of a piece of micanite wound with nickel wire as shown in Fig. 15·14 (*b*). This particular form of winding is adopted so that the outer portion of the mat shall act as a guard-ring to the central portion. This mat is insulated from the brass plates by asbestos. Above the hot plate is a compound plate made of ebonite of known thermal conductivity. The thickness of the central portion, B, of this plate is determined. The junctions of a manganin-constantan thermocouple are placed above and below the central portion of the ebonite, so that the temperature difference across the ebonite B may be determined accurately. The junctions are near to the centres of the faces of B. C_1 is a cold-water chamber placed on top of the ebonite. A rapid stream of water passes through this chamber, the interior of which is divided into channels so that the water flows as in Fig. 15·13 (*c*). Below the heating element is a square chamber, D, containing the liquid whose heat conductivity is to be measured. D is made by fitting two brass plates into a thin square ebonite frame. C_2 is a cold-water chamber similar to C_1. The stream of water through C_1 and C_2 is sufficiently rapid for its rise in temperature to be so small that the temperatures of the lower surface of the water and the upper surface of the ebonite are uniform.

Under these conditions, when electrical energy is dissipated in the heating unit, the temperature gradient in B and in D is large and may be measured with sufficient precision by means of calibrated thermocouples. The method of determining the temperature gradient in the ebonite has been indicated. For the determination of the temperature gradient in the liquid, a second manganin-constantan thermocouple is arranged as shown. The actual junctions are at X and Y, the wires being supported at a measured distance apart by fixing the manganin portions to an ivory pillar P_1 and supporting the constantan wire on a second pillar P_2.

The heating current is supplied from a large battery and the rate of supply of energy in the central portion of the heating coil determined by measuring the current through the wire and the voltage across the portion of the wire which is surrounded by the guard-ring.

To improve the thermal contact between the various surfaces the apparatus was clamped between wood supports and placed in melted wax. The whole was then placed in an exhausted vessel, and the air between the various surfaces was removed. Air was then admitted to the exhausted vessel, and when the wax was about to solidify the apparatus was removed. The crevices between the various surfaces were then filled with wax.

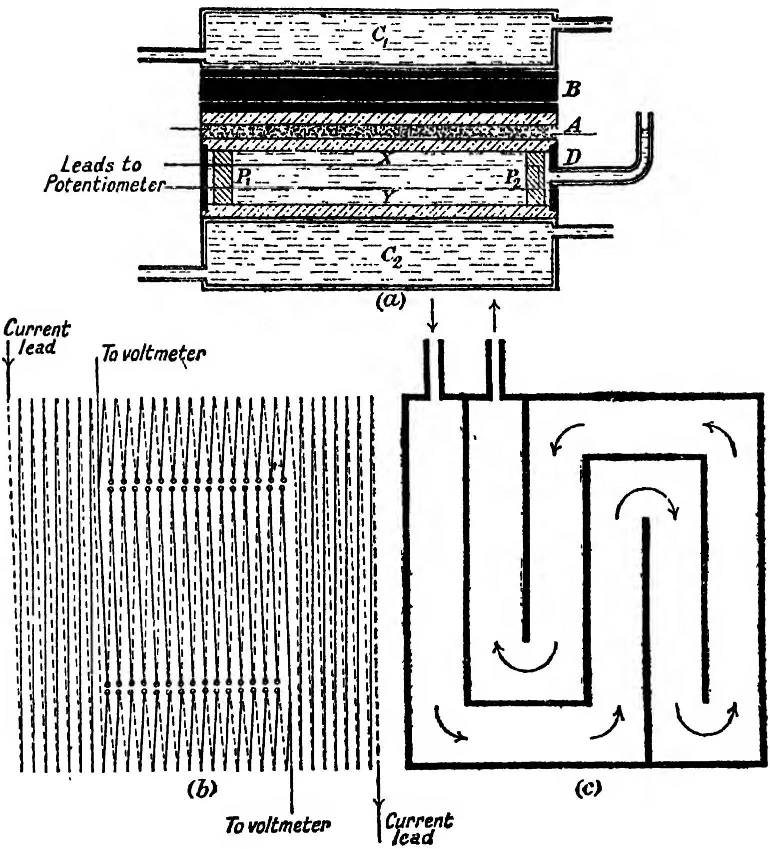


Fig. 15·14.—Apparatus for determining the Thermal Conductivity of Liquids.

By arranging the apparatus in this way convection currents in the liquid are avoided, and the guard-ring ensures that the flow of heat over those portions where the temperature gradient is measured shall be normal to the faces of the ebonite and the layer of liquid.

If W is the rate of supply of energy to the heating element in watts, A the central area of the mat, κ_1 the thermal conductivity of the ebonite, κ_2 that of the liquid, d_1 the thickness of the ebonite B, d_2 the distance between the junctions of the thermocouple in the liquid, θ_1 the tem-

perature drop across the ebonite B, and θ_2 that across the liquid of depth d_2, then, in the steady state,

$$\frac{W}{JA} = \kappa_1\frac{\theta_1}{d_1} + \kappa_2\frac{\theta_2}{d_2}.$$

From this equation κ_2 may be determined, since all other quantities in it are known or measurable.

The Thermal Conductivities of Gases.—The conductivity of a gas is very low and its measurement is again made difficult by the existence of convection currents. The following experiment is to illustrate the wide limits between which the conductivities of gases may vary. A platinum wire AB, Fig. 15·15, is suspended inside a wide glass tube. Its upper end is attached to a copper rod held in position by a cork C and connected to one pole of a battery. The wire is kept taut by a weight near B to which is attached a short metal rod dipping into mercury as shown. A copper wire passing through the lower cork D connects B to the other pole of the battery, the circuit containing a variable resistance and an ammeter as indicated. The current is adjusted until the wire glows. If a platinum wire is not available one of nickel may be used. The key K is then removed and coal-gas—or hydrogen—passed through the apparatus. The exit tube is immersed below the surface of water contained in a metal dish. If a small gas flame, or preferably a wire made hot by the passage of an electric current through it, is held over the water the gas may be disposed of without fear of an explosion. When the gas has been passing for several minutes the key, K, is closed, and although the ammeter indicates the passage of a current the wire no longer glows. This is partly because the conductivity of hydrogen is seven times that of air—some heat is lost by convection and by radiation.

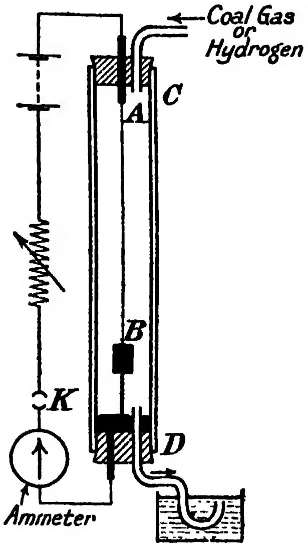


FIG. 15·15.

Convection Currents in Air.—A wax candle is attached to a piece of lead so that it may stand upright when placed in a shallow dish containing water—Fig. 15·16. The candle is lighted and a glass tube of the shape indicated placed over it, the water making an effective seal at the bottom of the tube. In a few seconds the flame is

extinguished. If, however, a metal T-piece is placed in the neck of the tube and the experiment repeated the candle continues to burn. If two glass rods, one moistened with strong hydrochloric acid and the other with ammonium hydrate, are held close together at A, white fumes of ammonium chloride will reveal that the air is entering the tube as indicated by the arrows.

The draught of a chimney is produced by convection currents. Similar currents in the atmosphere are responsible for the ***Trade Winds*** which blow with great regularity over certain portions of

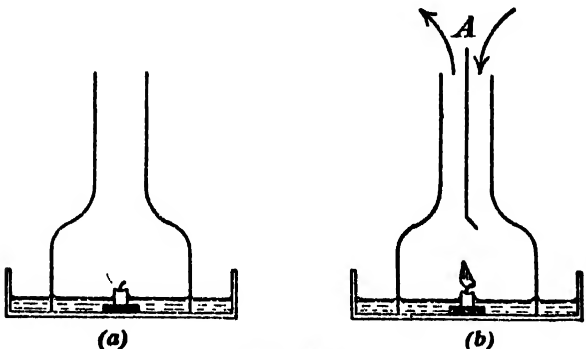


FIG. 15·16.

the earth's surface. They are produced by the cooler air flowing in from the north and south temperate zones to replace the hot air which is continuously moving upwards as a convection current in regions near the equator. The region where this hot air rises is the region of the equatorial calms. The rotation of the earth prevents these winds from following a course parallel to a line of longitude, since the velocity at the earth's surface becomes less as the latitude increases. Hence the wind in regions of higher latitude will lag behind that at the equator, appearing to come from the N.E. and S.E. in the northern and southern hemispheres, respectively.

Newton's Law of Cooling.—If a body is suspended in air and surrounded by a vessel whose walls are at a temperature lower than that of the body itself, Newton's law of cooling states that ***the rate of loss of heat*** [***in calories per unit time***] ***from the body at any instant is directly proportional to the excess of temperature of the body over that of its surroundings, if other conditions remain constant.*** Experiment has shown that this law is true only providing that the temperature difference between the body and the surroundings is not large.

To verify this law the apparatus shown in Fig. 15·17 may be

used. A copper sphere, about 4 cm. in diameter, is suspended by three fine wires. The bulb of a mercury-in-glass thermometer or one junction of a thermocouple, is inserted in a hole drilled in the sphere. Good thermal contact between the bulb and the thermometer is obtained by filling the hole with mercury. The sphere is raised to the desired temperature by heating it on a hot sand bath. The latter is then removed and the sphere suspended within an enclosure whose walls are at a known temperature : a cooling curve is constructed from observations on the temperature of the bulb at different times—cf. Fig. 15·18 (*a*).

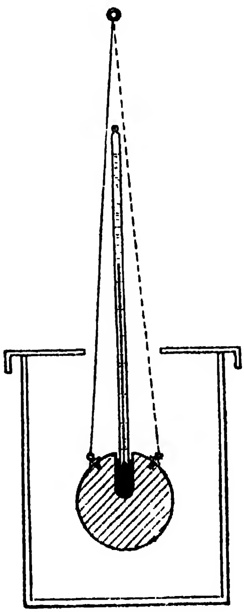


FIG. 15·17.

The rate at which the temperature of the body at different times is changing is ascertained in the manner indicated on p. 212. A graph showing the relation between the rate of cooling and the excess temperature is then plotted—cf. Fig. 15·18 (*b*). But further considerations are necessary before the shape of this graph will enable us to find out whether or not Newton's law of cooling is true. This law refers to a rate of loss of heat [number of calories lost per unit time], whereas we have only dealt with the rate of cooling [change in temperature per unit time]. Suppose the temperature of the sphere changes by an amount $\delta\theta$ in time δt. The mean rate of cooling during this interval is $-\frac{\delta\theta}{\delta t}$, and in this time the heat lost is $-(\mathrm{MS}+ms)\delta\theta$, where M is the mass of the copper, m that of the mercury, S the specific heat of copper, and s that of mercury. The negative sign occurs since $\delta\theta$ is essentially negative and the heat lost is a positive quantity. The rate of loss of heat is

$$-(\mathrm{MS}+ms)\frac{\delta\theta}{\delta t}.$$

From this we see that the rate of cooling is proportional to the rate at which heat is lost, only if the thermal capacity of the sphere, etc., is constant. These are constant if the specific heats S and s are constant. This is one reason why a copper sphere was selected : another is that it is undesirable to use liquids in this experiment on account of evaporation losses. Hence a straight-line relation between the rate of cooling and the excess temperature is, in this

instance, a verification of the validity of Newton's law of cooling over the range of temperature investigated.

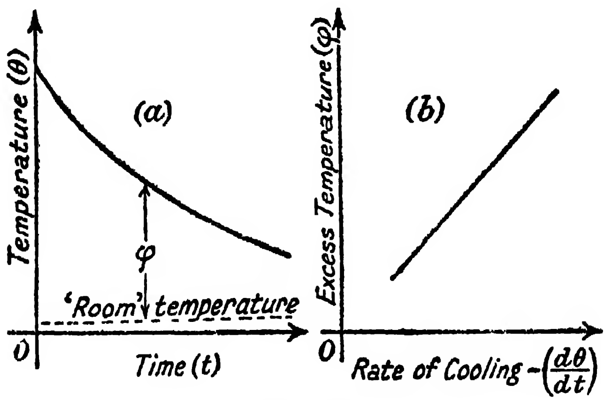


FIG. 15·18.

It is interesting to note that NEWTON actually performed this experiment in a current of air, i.e. a draught. Thus the convection was ' forced ' and the radiation losses comparatively small. Under these conditions the law is valid for quite large temperature differences.

Example.—A copper ball cools from 62° C. to 50° C. in ten minutes, and to 42° C. in the next ten minutes. Calculate its temperature at the end of the next ten minutes.

Let t° C. be the room temperature. The average rate of cooling in the first ten minutes is 1·2° per minute and this may be taken as the rate of cooling at the mean temperature 56° C. Since this is proportional to the excess temperature (the thermal capacity of the ball being constant), we have

$$1{\cdot}2 = \kappa(56 - t),$$

where κ is a constant. Similarly for the next ten minutes

$$0{\cdot}8 = \kappa(46 - t).$$

By division we have

$$1{\cdot}5 = (56 - t) \div (46 - t).$$

Therefore $\quad t = 26{\cdot}0°$ C. Whence $\kappa = 0{\cdot}04$ min.$^{-1}$.

Let θ be the temperature at the end of the next ten minutes. Then the rate of cooling is $\dfrac{(42 - \theta)}{10}$, while the mean temperature is $0{\cdot}5(42 + \theta)$. Hence by Newton's Law,

$$0{\cdot}1[42 - \theta] = 0{\cdot}04[0{\cdot}5(42 + \theta) - 26].$$

Hence $\theta = 36{\cdot}7°$ C.

EXAMPLES XV

1.—The temperature difference between two opposite faces of a metal plate is 40·6° C.; each face measures 30·4 cm. × 25·6 cm. If the thickness of the plate is 4·82 cm., calculate the conductivity of the metal if the heat passing through the plate is sufficient to melt 582 gm. of ice per minute. [l = 80 cal. gm.$^{-1}$.]

2.—Define the term *thermal conductivity* and describe how you would proceed to measure the thermal conductivity of a badly conducting solid if it were available in a form suitable for the method you select.

3.—What is meant by the statement that the thermal conductivity of iron is 0·15 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C.? Calculate the amount of heat which will flow per minute through a sheet of iron 1 metre square and 4·5 mm. thick if one face is at 100° C. and the other at 110° C.

4.—The glass windows of a room have a total area of 10 square metres and the glass is 3 mm. thick. Calculate the rate at which heat escapes from the room by conduction when the inside surfaces of the windows are at 20° C. and the outside surfaces are at − 5° C. [κ = 0·002 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C.]

5.—Describe and explain a method of measuring the thermal conductivity of glass, the glass being supplied in the form of a tube.

6.—Describe how the thermal conductivity of a badly conducting solid available in the form of two rectangular blocks of the same size may be determined, indicating clearly how the conductivity is calculated from the observations.

7.—Describe and explain the principle of a miner's safety lamp and state the conditions under which such a lamp may become dangerous.

8.—A flat heating coil, in which energy is dissipated at the rate of 20 watts, fills the space between two identical cylindrical discs 20 cm. in diameter and 0·2 cm. thick and the whole is suspended in air. The mean temperature difference between the faces of each disc when steady conditions have been obtained is 10° C. Calculate a value for the thermal conductivity of the material of the discs.

9.—In an experiment with Searle's apparatus to determine the thermal conductivity of copper the following observations were made. t_1 = 73·9° C., t_2 = 50·4° C., t_3 = 17·50° C., t_4 = 14·21° C. Mass of water flowing in 36·9 secs. = 500 gm. Distance between t_1 and t_2 = 10·0 cm., diameter of bar 5·07 cm. Calculate the thermal conductivity of copper.

10.—A flat heating coil completely fills the space between two sheets of ebonite each 2 mm. thick. The whole of the above is then inserted between two plates of glass, each 8 mm. thick, there being good thermal contact between adjacent glass and ebonite surfaces. The cross section of the above composite block in any plane parallel to that of the heating coil is 1000 cm.^{2}. The temperature of each ebonite face in contact with the heating coil is 50° C., the outer faces of the glass are maintained at 0° C. The thermal conductivities of ebonite and of glass are 0·4 × 10^{-3} and 2 × 10^{-3} cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C., respectively. At what rate is energy being dissipated in the heating coil?

11.—State *Newton's law of cooling*. Describe and explain how you would test the validity of this law experimentally.

A copper sphere is heated and allowed to cool while suspended in an enclosure whose walls are maintained at a constant temperature. The sphere is found to cool from 92° C. to 80° C. in 10 minutes and from 80° C. to 70° C. in the following 10 minutes. Calculate a value for its temperature at the end of the next 10 minutes.

CHAPTER XVI

THE TRANSMISSION OF HEAT—RADIATION

Preliminary Remarks.—Heat can be transferred by conduction and convection only through a material medium, solid or fluid in the former instance, fluid alone in the latter, but the fact that we receive heat from the sun provides ample evidence that one body may heat another even though the two bodies are separated by a space devoid of ordinary matter. The process by which this occurs is known as ***radiation,*** and while in course of transfer the heat energy takes a form spoken of as ***radiant energy.***

The transfer of heat by radiation is not limited to empty space, however, for some at least of the radiant energy emitted by the sun reaches the surface of the earth in spite of the layer of air covering it. Hence radiant energy can pass through a gas. Moreover, it is an everyday experience that it can pass through glass, and experiments, to be described later, show that it passes even better through rock salt and carbon disulphide. Finally, it may be noted that some substances opaque to visible light allow radiated heat to pass through them: ebonite is one of these, a solution of iodine in carbon disulphide another. In order to understand the processes at work let us consider the following analogy.

Experiment.—Two similar tuning-forks are mounted on resonance boxes so that they lie in the same plane at a short distance from each other. One of the forks is bowed strongly: the waves emitted travel through the air and impinge upon the second fork. If the prongs of the first fork are held so that they no longer vibrate, a note from the second fork will be heard, although it was silent originally. This is an example of the radiation of sound energy and its reception by a body of the same natural frequency.[1]

Since matter consists of molecules moving in all directions at random (liquids and gases) or oscillating about some mean position (solids), we have to liken matter to a swarm of tuning-forks.

[1] Every tuning-fork has a definite period of vibration—say T sec. The reciprocal of T, which is numerically equal to the number of vibrations per second is the frequency, f, of the fork. Thus

$$= \frac{1}{T} \text{ cycle. sec.}^{-1}.$$

In general, whatever is the nature of the radiant energy incident upon a body, some of the molecules present in that body will be able to act as receivers for that part of the radiation which they themselves would emit had they been stimulated. If a set of tuning-forks emit radiations which fall upon another similar set of forks, these will continue to absorb energy until they themselves are emitting energy at a rate equal to that at which it is being received. From this point of view the thermal equilibrium of a body is not one of rest but of vigorous activity, for a body at a constant temperature is one in which there is perfect compensation between the heat it absorbs and the heat it radiates.

The Early History of Radiant Energy.—The history of radiant energy dates from the time of FRANCIS BACON. For centuries before, men had known how to use burning mirrors to concentrate the sun's rays to a focus and thereby kindle a fire. Bacon suggested the use of burning glasses to concentrate 'the heat which is not glowing or luminous, but such as the heat of iron or stone which has been heated but not ignited, or the heat of boiling water.'

Diathermancy.—When white light is incident upon a body, in general, part will be transmitted, part reflected, and the remainder absorbed. For our present purpose it is sufficient to know that white light is a mixture of different colours and that each colour is characterized by a certain frequency of vibration. When colours of particular frequencies are absorbed those which remain cannot produce white light—we find that the transmitted and reflected light is coloured. The body is *opaque* to the visible rays which it absorbs, and *transparent* to those it transmits. In the study of radiant energy substances are found, in general, to behave in a similar way. Substances transmitting heat radiations of particular frequencies are said to be *diathermanous* with respect to those radiations: similarly, a substance absorbing such radiations is *adiathermanous* with respect to them. The two words diathermanous and adiathermanous correspond to the terms transparent and opaque in optics.

Instruments used in Detecting Heat Radiation.—When radiant energy is incident upon an absorbing material the temperature of the latter increases. If the effect of this rise in temperature, which is often small, may be amplified, we have a detector of radiant energy. No known substance absorbs all incident radiation: lamp-black, however, absorbs more than 90 per cent., irrespective of the particular source from which the radiation may come. The early workers in this field used differential air thermometers, one bulb being covered with lamp-black. These have been superseded by electrical instruments, based on the fact that an electric current flows continuously in a circuit consisting of two dissimilar metals

when the junctions of the metals are at different temperatures. The effect is very marked when the metals are antimony and bismuth, the current flowing from the antimony to the bismuth through the cold junction. A set of such antimony-bismuth junctions coated with lamp-black and arranged so that the separate effects are additive becomes a sensitive detector of heat radiation and is termed a ***thermopile.*** Fig. 16·1 (*a*) is a diagrammatic representation of the arrangement of the small bars of metal in a thermopile. They are insulated along the greater part of their lengths by mica and embedded in pitch, or other insulating material, with their junctions projecting at the two ends. Fig. 16·1 (*b*) is an end-on view of a thermopile consisting of twenty-five junctions. One set of junctions is polished and covered by a metal cap: the

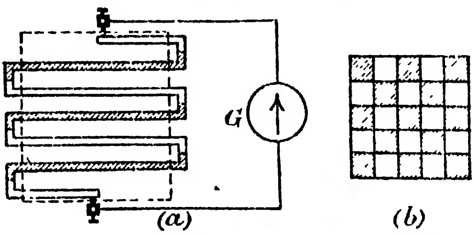


FIG. 16·1.—A Thermopile—its construction and use.

other set is coated with lamp-black [to absorb heat radiations more readily] and exposed to the source to be examined. A metal cone screens the blackened face from all radiations except those lying within the angle of the cone. This instrument was invented by NOBILI in 1829 and subsequently improved by himself and by MELLONI. Of still more recent date is the bolometer, which is essentially a strip of very thin platinum foil coated with lamp-black. When exposed to heat radiations its electrical resistance increases in consequence of the rise in temperature experienced. The change in resistance is measured by some form of Wheatstone bridge.

To measure the energy associated with a small region of a spectrum the alternate junctions must be arranged in a straight line one above the other. We then have a ***linear thermopile.*** Fig. 16·2 (*a*) is a diagram showing the construction of such a thermocouple. Silver wire, 0·03 mm. in diameter, and bismuth wire, 0·1 mm. in diameter, are used to form the individual thermocouples. Since bismuth wire is very brittle, short pieces are employed—they must be sufficiently long, however, to ensure that the temperature of the 'cold junction' remains constant. Tin is used to join the wires to small copper discs. Fig. 16·2 (*b*) shows

how the thermocouples are assembled. The 'hot junctions' lie in a straight line and their surfaces are blackened. The whole is enclosed in a metal case in front of which is a narrow slit, so that for a given setting of the apparatus only radiation in a small region of the spectrum falls on the 'hot junction.'

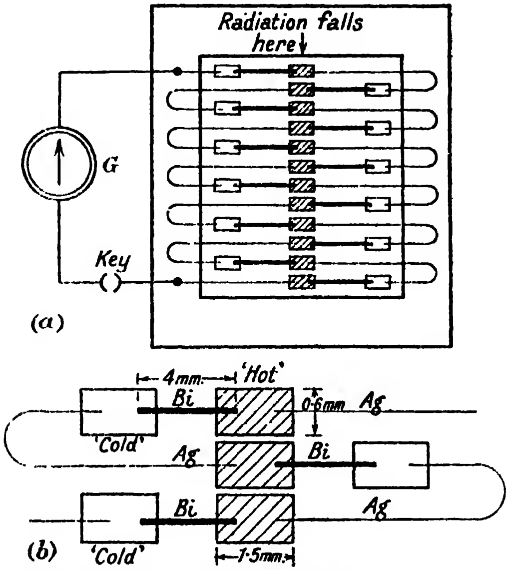


FIG. 16·2.—Construction of a Linear Thermopile. [Not to scale.]

The sensitivity of thermopiles has been considerably increased in recent years by making the mass of the instrument small and enclosing it in a vacuum to render negligible the loss of heat by conduction and convection from its surface.

The Rectilinear Propagation of Radiant Energy.—That radiant energy travels in straight lines may be demonstrated by arranging three horizontal narrow tubes each about 5 cm. long in a straight line between a hot body and a thermopile. A galvanometer suitably connected to the thermopile indicates that it is receiving energy unless the collinear arrangement is destroyed by displacing one of the tubes.

The Inverse Square Law for Radiant Energy.—If a small element of area is constructed so that it is perpendicular to the direction of flow of radiant energy at a point, the linear dimensions of the source being small, the amount of energy passing through that element per second divided by the area of the element is termed the ***intensity of the radiation*** at that point. The ***inverse square law*** states that the intensity of radiation at a point is inversely

proportional to the square of the distance of that point from the source. This statement may be verified as follows :—A thermopile, T, Fig. 16·3, connected to a galvanometer, G, is placed in front of a large tank, M, filled with boiling water (or otherwise maintained at a steady temperature). The surface of the thermopile is directed towards the tank and the deflexion of the galvanometer recorded. It will be found that as long as the thermopile is moved along a normal to the surface of M, this deflexion remains constant providing the generators of the cone of the instrument

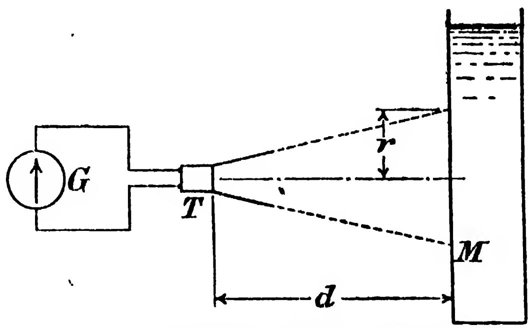


FIG. 16·3.—Inverse Square Law for Radiant Energy.

do not pass beyond the confines of M. For if d and r are, respectively, the distance of M from T, and the radius of the circle on M, from which the thermopile receives radiation, and suffixes denote corresponding conditions, $\frac{r_1}{d_1} = \frac{r_2}{d_2}$. Moreover, the areas of the circles are πr_1^2 and πr_2^2. Since the galvanometer gives a constant deflexion it follows that the total energy received by it is the same in each instance. We therefore have

$$\frac{\text{Intensity of radiation from 1 cm.}^2 \text{ of surface at } d_1}{\text{Intensity of radiation from 1 cm.}^2 \text{ of surface at } d_2} = \frac{d_2^2}{d_1^2},$$

since experiment shows that the terms formed by the cross-multiplication of the above fractions are equal.

The Reflexion of Radiant Energy.—Fig. 16·4 is typical of an arrangement whereby the laws of reflected radiant energy may be established. Two brass tubes (15 cm. × 0·2 cm.), LM and PQ, are placed in the same horizontal plane before a piece of polished metal sheet B capable of rotation about a vertical axis passing through the intersection of imaginary vertical planes drawn along the axes of the tubes. A white-hot ball, or an arc lamp, is placed at A and the thermopile at C to receive any radiation passing down the tube PQ. The metal B is rotated until the deflexion of the galvanometer, G, shows that C is receiving radiant energy.

Since it is difficult to measure the angles at B directly, the thermopile is removed and the eye directed along the tube QP, care being taken to hold a piece of smoked glass near Q to protect the eye from ultra-violet rays when an arc is employed. A clear image

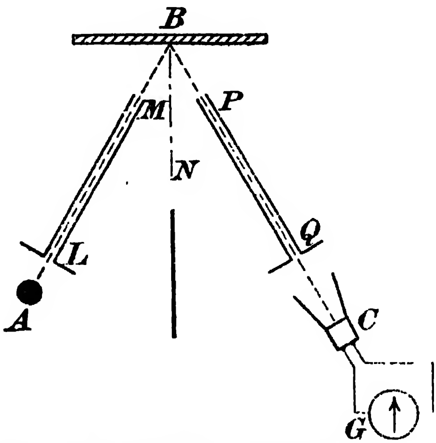


FIG. 16·4.—Reflexion of Radiant Energy.

of the source will be seen showing that the laws for the reflexion of radiant energy are the same as those governing the reflexion of light [cf. p. 368].

Further Experiments on the Reflexion of Radiant Energy.—If an arc lamp is placed at the focus, B, of a concave mirror,

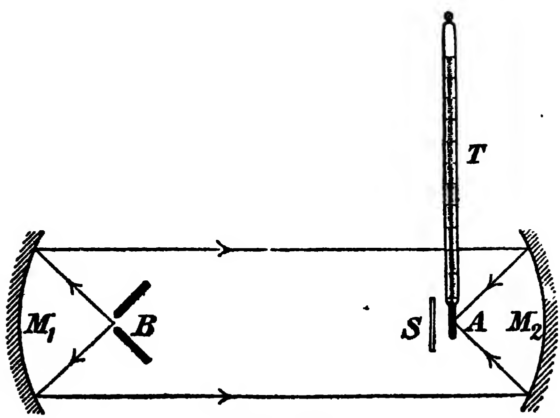


FIG. 16·5.

M_1, Fig. 16·5, radiant energy is reflected from the mirror as a parallel beam [cf. p. 382]. Suppose that this falls on a second

concave mirror M_2 having the blackened bulb of a thermometer, T, at its focus A. A cardboard screen S protects this bulb from direct radiation from B. If the mirrors have a focal length about 12 cm. and are placed one metre apart, the rise in temperature at A is about 4° C. This experiment is a verification of the fact that radiant energy is propagated according to the laws of geometrical optics.

The Refraction of Radiant Energy.—To verify the fact that radiant energy may be refracted, an image of a slit S, illuminated by an arc lamp, A, Fig. 16·6, is produced by a converging lens, L_2, on the front surface of a thermopile T at T_1. If the latter is connected to a galvanometer the deflexion shown by this instrument proves that thermal energy is incident upon the surface of the thermopile. When a hollow glass prism, P, containing carbon disulphide, is introduced into the path of the light emerging from the

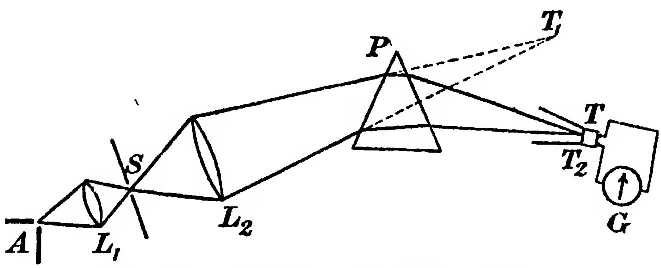


FIG. 16·6.—Refraction of Radiant Energy.

lens the galvanometer deflexion soon becomes zero and a spectrum may be obtained on a screen held in a suitable position. The prism P is then rotated until the deflexion of the light passing through it is a minimum [cf. p. 409]. The thermopile, T, is then placed at T_2 to receive this spectrum. It is advisable to cover the exposed surface of the thermopile with a cardboard having a narrow slit parallel to the refracting edge of the prism so that only one colour falls on T at one time. If the violet rays, i.e. the rays of short wave-length, are first allowed to pass through the slit a small deflexion will be shown by the galvanometer. This deflexion increases rapidly as the green, yellow, and red rays are in turn allowed to reach T. If the region beyond the red rays is explored in this way it is found that the deflexion of G continues to increase for some time before again becoming zero, a fact showing that the region immediately beyond the red end of the spectrum is rich in heat rays. These are termed ***infra-red rays.***

The Distribution of Energy in a Heat Spectrum.—By means of an apparatus similar to that shown in Fig. 16·6

[cf. also Fig. 23·18], the distribution of energy in the spectrum of the heat radiation from a hot body may be investigated. The essential modifications are that the thermopile should be a linear one, and that the materials of the prism and lenses should be diathermanous to heat rays. The galvanometer deflexion will be directly proportional to the energy received by the thermopile per unit time, i.e. to the energy in a short region of the spectrum. If the deflexions are plotted against the wave-lengths [1] of the heat radiations corresponding to the centres of each such short region, curves similar to those shown in Fig. 16·9 [cf. p. 337], will be obtained. These indicate that as the temperature of the source increases, the maximum on the curve shifts towards the region of shorter wave-length, i.e. the region of higher frequency.

Early Experiments on the Amounts of Heat lost per Unit Time from Equal Areas of Different Surfaces under Identical Conditions.—To compare the rates of emission of radiant energy from equal areas of different substances at the same temperature LESLIE devised the following experiment. A metal cube, side about 10 cm., was filled with boiling water [if the experiment were being repeated, electrical heating would be employed to keep the water boiling], and placed in front of a thermopile connected to a galvanometer. Three of the side faces of this cube were covered with the materials under investigation—say lamp-black, varnish, and paper, while the fourth was highly polished. The thermopile was at such a distance from the cube that only heat from the surface under examination was received by it. The deflexion of the galvanometer was proportional to the rate at which energy was received by the thermopile. The radiation from each different face of the cube was examined in turn, care being taken to keep each face at a fixed distance from the thermopile. Leslie found that lamp-black was the most efficient emitter of radiant energy, while polished metal surfaces were very inefficient.

With the aid of this apparatus it may be shown that aluminium paint is a poor emitter of heat; hence, as far as the emission of radiant energy is concerned, it is most disadvantageous to coat hot-water pipes with aluminium paint.

Experimental Investigation of the Diathermancy of Different Bodies.—MELLONI compared the diathermancy of various substances in the following way: A screen, S, Fig. 16·7, having a circular opening, was arranged as indicated between a source of radiant energy, A, and a thermopile, T, connected to a galva-

[1] If c is the velocity with which radiant energy travels, f cycle. sec.$^{-1}$ the frequency, then the wave-length λ is such that $c = f\lambda$.

nometer, G. A was a steam chamber, and the opening in S was such that when A and T were in position only radiant energy from the near side of A was received by T. The deflexion of G having been noted, a piece of glass plate was placed at P. The deflexion was considerably reduced. The ratio of the galvanometer deflexions in the two instances measured the diathermancy of the particular piece of material used. When a second piece of glass similar to the first was placed alongside P the deflexion was not much reduced, a fact which showed that a substance is very diathermanous towards radiation transmitted through some of the same material.

From such experiments as these it was found that rock salt was the most diathermanous substance investigated—that is why the lenses and prisms used in experiments on radiant energy should preferably be made of this substance.

Liquids may also be examined in this way. Since they have to be contained in a glass cell a blank experiment is first performed with the cell empty.

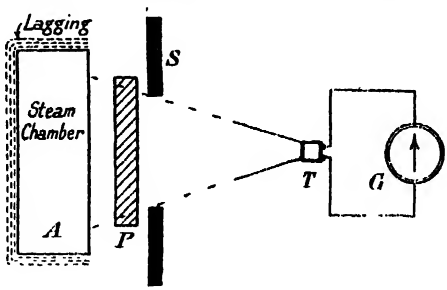


FIG. 16·7.—Apparatus for comparing the Diathermancy of Different Bodies [after Melloni].

Water is less diathermanous than glass, but even as ice it transmits heat rays without melting. This may be shown by filling a watch-glass with water and placing it on solid carbon dioxide (— 80° C.), or on a freezing mixture. The plano-convex lens so formed may then be used to form an image of an arc lamp on a thermopile when a galvanometer connected to it will show an increased deflexion.

Ebonite and a solution of iodine in carbon disulphide are diathermanous with respect to heat waves although they are opaque to visible radiations.

Melloni found that the diathermancy of a body increased in general as the temperature of the source increased. As an example of this we may cite the instance of glass, which allows heat rays from the sun to pass through it without becoming warm. The outer layers of the sun are estimated to be at 6000° C. [JEANS estimates that inside the sun the temperature may exceed one million degrees.] On the other hand, glass is used as a fire-screen because it is adiathermanous with respect to heat rays from a source at a relatively low temperature, about 1,200° C. It is for

these reasons that glass is employed in greenhouses. Heat from the sun may pass through, but the heat from the materials inside cannot.

These facts may be illustrated in the laboratory by suspending a piece of copper sheet in a vertical position, and heating it in a bunsen flame. If a piece of glass is placed between the hot copper [about 400° C.] and a thermopile, the galvanometer to which this is connected shows that the thermopile is not receiving much radiant energy—glass does not transmit readily the heat energy from such a source. If the hot copper is replaced by an Argand burner similar results will be obtained. When, however, an arc lamp is used as the source of radiant energy, the fraction of the energy transmitted is much greater.

Prevost's Theory of Exchanges.—The early workers who endeavoured to ascertain the nature of radiant energy were confronted by the following difficulty. It is well known that a hot body radiates heat to those bodies that are cooler than itself, but does it also radiate energy when it is surrounded by bodies at a temperature equal to or greater than its own, i.e. does the radiation from a given body at a given temperature depend on the objects surrounding it, or is it independent of them? According to the theory of exchanges due to PREVOST, who termed it 'A Theory of a Movable Equilibrium of Temperature', it is maintained that bodies at all temperatures are continuously radiating energy to each other, those at a constant temperature receiving as much energy in a given time as they emit. To see how he arrived at such a conclusion let us consider, with Prevost, a number of bodies initially at different temperatures in an enclosure whose walls are impervious to heat and which contains no source of thermal energy. Ultimately these bodies will acquire a uniform temperature and be in equilibrium with each other and the walls of the enclosure. This condition is reached by a process in which energy is both absorbed and emitted, and not by one in which the hot bodies emit energy and the colder ones receive it. It is independent of the size, shape, and position of the objects with respect to the walls. Moreover, this theory asserts that this mutual process of the simultaneous emission and absorption does not come to an end when thermal equilibrium is attained, but that there is a continuous exchange of energy between the bodies themselves and between each body and the walls of the enclosure, although the total energy in each body remains constant.

Now if one of the bodies were withdrawn from the above enclosure and placed in another whose walls, and the objects in it, were in thermal equilibrium with one another but at a temperature higher than its own, heat would be radiated from the walls, etc.,

to that body. The body thus introduced was not capable of acting directly on the walls of the enclosure and the objects therein which were at a distance from it, i.e. the cooler body could not have caused the walls, etc., to emit radiant energy to it. The mutual processes of emission and absorption could not therefore have ceased when thermal equilibrium had been attained inside the enclosure. The theory of exchanges is based on arguments similar to the above.

When a body is placed in an enclosure whose walls are at a temperature equal to its own, the temperature of the object remains constant because the heat it receives from the walls is exactly balanced by the heat it gives to them. For if the object became hotter than they, its rate of supply of thermal energy to the walls would at once become more copious and thermal equilibrium would soon be re-established.

In this connexion it is well to remind ourselves that a thermometer suspended in a room may not indicate the temperature of the air in its immediate neighbourhood even if that temperature is steady and the thermometer has not just been placed in position: for its indication will depend on the nature of the radiations which its sensitive part is receiving from surrounding bodies, if it is able to absorb them.

Further Evidence in Support of the Theory of Exchanges.—Let us assume the validity of the theory and see whether some of the consequences ensuing from it are in accord with experimental facts. Suppose that two mercury thermometers, identical in all respects, except that the bulb of one is blackened while the other is enclosed by a silver thimble in good thermal contact with it, are placed in an enclosure whose walls are maintained at a constant temperature. The final indications of the thermometers are identical. All the radiant energy falling on the bulb of the first thermomoter is absorbed, whereas that falling on the silver is mostly reflected. Since the temperatures recorded are the same, however, it follows that the bulb which is blackened must be emitting a supply of energy equal to the amount it receives when thermal equilibrium is reached, whereas the bulb of the other thermometer only emits a correspondingly small amount, but again equal to that which it receives.

The radiation from a reflecting metallic surface ought, therefore, if the theory is true to be much less than that from a blackened one at the same temperature. Leslie proved experimentally that surfaces which reflect radiant energy copiously only emit a small amount at the same temperature.

An Important Theorem.—Suppose that I is the amount of radiant energy received per unit time by a body. Let A, R, and

T denote the amounts absorbed, reflected, and transmitted, respectively. Then

$$I = A + R + T.$$

If A is large, it follows that R and T are small, i.e. a good absorber is a poor reflector. Similarly, a good reflector is a poor absorber.

The Radiant Energy from Heated Substances.—Suppose that a thin plate of rock salt is suspended in a temperature enclosure of the type already considered. The temperature of the plate finally assumes a constant value, when it radiates as much energy per unit time as it absorbs in that time. Since rock salt is diathermanous [transparent to heat radiations], the rate of emission of radiation from it will be small. Moreover, since a thick plate of rock salt will absorb more than a thin one in the same time, it will also radiate more. Balfour Stewart verified these deductions experimentally.

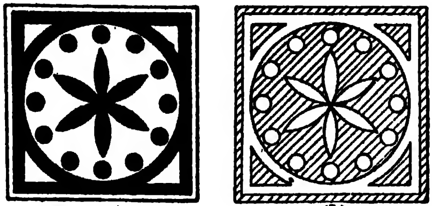
Suppose that the plate is made of glass and the temperature of the enclosure is not high—say 400° C. Since glass is extremely adiathermanous [opaque to heat rays], either a thick or a thin plate will absorb nearly all the heat energy incident upon it. The radiation from such plates will therefore be independent of their thickness—in fact, the rate of emission from either is the same practically as if they were coated with lamp-black.

Experiments such as the above show that the surface of a body is not necessarily the source of the heat radiations.

The Extension of the Theory to Bodies exchanging Radiations at Different Temperatures.—When radiations of all wave-lengths fall on an object and are absorbed by it, it does not follow that radiations of all wave-lengths are emitted except when the object is in thermal equilibrium with the surrounding objects; if the body is cold the greater portion of the radiation emitted will have long wave-lengths, but the total energy emitted will be equal to that absorbed when conditions are steady. If the substance is heated the proportion of radiation having short wave-lengths increases—ultimately visible rays are emitted. Now at ordinary temperatures the black portions of the design on the china used below absorb more (in fact nearly all) of the incident radiation than do the red; they therefore emit more. Similarly, the red portions absorb and therefore emit more energy as radiation than do the white in the pattern: if this were not so the black portions would be much hotter than the red and these much hotter than the white. When the temperature of the china is raised a point is reached when the black portions emit a copious supply of visible radiation, the red less, and the white least of all.

This apparent reversal of a black and white pattern on heating

is indicated in Fig. 16·8. In connexion with this, it must be remembered that so long as we are dealing with heat rays, the fact that one part of a surface is a better radiator of energy than another, may only be ascertained with the aid of an instrument which will detect such radiations—the eye fails utterly. It is for this reason that the white portions in Fig. 16·8 (*a*) appear brighter than the dark ones to an observer. If the temperature is so high, however, that visible rays are emitted, ($\lambda < 0{\cdot}7\ \mu$), the difference in the radiating powers of two surfaces is then apparent—in fact, we have a ***visible*** proof that a blackened object radiates more than a white one at the same temperature.



(*a*) (*b*)

FIG. 16·8.—Apparent Reversal of a Pattern on Heating.

Experiment i.—Obtain a piece of white china having a design in red and black. Heat it in a strong blowpipe flame until it is incandescent. The parts which were white originally, now appear darkest and vice-versa, while the portions which were red at ordinary temperatures now glow, but less vividly than those portions originally black.

Experiment ii.—Mark with ordinary black ink [or, better, a paste of ink and iron oxide] a cross on a piece of platinum foil and heat it in a blowpipe flame. The ink is converted to oxide of iron which glows more vividly than the rest of the foil when heated. In consequence of this greater emission from the oxide, the foil, when examined on the reverse side, will appear darker where the cross has been drawn on the other side.

Experiment iii.—Heat a rod of glass which is blue at ordinary temperatures. When it becomes incandescent it appears a very bright red. A piece which is red at ordinary temperatures shines less brightly than the above piece of glass when similarly treated. This is because blue glass absorbs red light in preference to blue, so that when it becomes incandescent red light is emitted more freely than the blue. On the other hand, the red glass absorbs blue light and so emits blue light more copiously than red when it is heated. A piece of transparent glass at the same temperature only gives out a faint light, since it absorbs very little of the incident radiation. If one heats a piece of yellow glass (absorbs blue) then it appears blue when the heating is effected in a darkened room.

'The Radiation of Cold.'—The following experiment, usually attributed to PICTET, although it was originally carried out by PORTA, is important in that it was probably the means whereby PREVOST of Geneva was led to propose his theory of exchanges which, as seen above, has played a considerable rôle in the study

of radiant energy. The experiment is practically the reverse of that described on p. 326. Pictet placed a lump of ice [or a freezing mixture] at the focus of one of the concave mirrors and the bulb of a thermometer at the other, the screen S, Fig. 16·5, remaining in position. The temperature of the thermometer fell.

The above experiment was at one time quoted as a proof that 'cold' was radiated from the ice. The modern explanation is as follows :—When the bulb of the thermometer is at room temperature it has acquired that steady state in virtue of the fact that it is receiving heat from all objects round it at the same rate as it itself is emitting radiation to them. When the block of ice was placed in position it acted as a screen protecting A from some of the radiation otherwise incident upon it, and emitted less radiation per unit time itself. A was therefore emitting more radiant energy to the ice than it received from it and cooled in consequence, until the rate of emission was again equal to that of absorption.

The Loss of Heat from Bodies by Radiation—Stefan's Law.—When a body cools in air the loss of heat from it takes place under rather complicated conditions, for the loss of heat depends on the processes of radiation, conduction and convection. Dulong and Petit, about 1817, carried out a series of researches in which they attempted to eliminate effects due to the two latter processes.

A mercury-in-glass thermometer was heated to 300° C. and then placed with its bulb at the centre of a copper sphere immersed in a water bath at constant temperature. The air pressure in the globe was reduced to about 2 mm. of mercury, and the rate of cooling of the thermometer observed. Their results were embodied in an empirical formula which, for many years, was thought to represent the rate at which a body emits heat radiation at a given temperature.

In 1879 Stefan suggested that the rate of emission of radiation from a body was proportional to the fourth power of its absolute temperature. Stefan was led to make this statement after a careful examination of some results published by Tyndall. This investigator found that at 1,200° C. the rate of emission of radiation from a platinum wire was 11·7 times the rate of emission at 525° C. Now

$$\left(\frac{1200 + 273}{525 + 273}\right)^4 = 11{\cdot}6.$$

[Callendar points out that Tyndall estimated the temperature of the wire from its colour: the above agreement is therefore fortuitous—the temperatures may be wrongly estimated by 100° C.!]

Stefan then examined the work of Dulong and Petit and found

that, if a correction for the residual gas in the apparatus was applied, their results were in accord with the fourth power law.

It must be borne in mind that this law, first suggested by Stefan and subsequently established by BOLTZMANN from theoretical considerations, which states that ***the total radiation emitted per unit time from unit area of a black body is proportional to the fourth power of its absolute temperature,*** does not mean that the rate at which a body cools, by losing energy in the form of radiation, is proportional to the fourth power of the absolute temperature, for the above rate also depends on the temperature of the enclosure in which it is situated. To obtain an expression for the rate of fall in temperature of a black body, mass m, specific heat s, in terms of its temperature $T°$ K and $T_0°$ K, the temperature of the enclosure, it is necessary to discover the rate at which the body receives heat energy from the walls. Let the rate of emission of heat energy from the black body be κT^4, where κ is a constant. Then when this body is at a temperature T_0 it will emit radiant energy at a rate κT_0^4; since its temperature will remain T_0, however, it must be receiving heat energy at a rate κT_0^4 and this will come from the walls of the enclosure. Consequently κT_0^4 is the rate at which the walls emit energy to the black body irrespectively of the temperature of that body; i.e. when the black body is at temperature T, the energy actually lost per unit time will be $\kappa T^4 - \kappa T_0^4 = \kappa(T^4 - T_0^4)$. Hence

$$ms[\text{Fall in temperature per unit time}] = \kappa(T^4 - T_0^4).$$

It is easily shown that in the particular instance when the temperature difference between the hot body and its surroundings is small, that the rate at which the hot body loses heat in the form of radiation [it must be suspended in an exhausted chamber], is directly proportional to the temperature excess.

For let $T = T_0 + \tau$, where τ is small. Then

$$\kappa[(T_0 + \tau)^4 - T_0^4] = \kappa T_0^4\left[\left(1 + \frac{\tau}{T_0}\right)^4 - 1\right]$$

$$= \kappa T_0^4\left[\left\{1 + 4\frac{\tau}{T_0} + 6\left(\frac{\tau}{T_0}\right)^2 + \text{terms in higher powers of}\left(\frac{\tau}{T_0}\right)\right\} - 1\right].$$

Since τ is small we may neglect $\left(\frac{\tau}{T_0}\right)^2$, and all its higher powers, in comparison with $\frac{\tau}{T_0}$, so that the above expression becomes $4\kappa\tau T_0^3$, i.e. the rate of cooling, under the conditions here stipulated, is directly proportional to the difference in temperature τ.

[It must be pointed out, however, that the above argument is not a theoretical proof of the validity of Newton's law of cooling,

for, as the sequel will show, this applies to the rate at which a body loses heat under very different conditions from those postulated here.]

'Black Body' Radiation.—The ideal black body is one which absorbs completely all the radiations incident upon it, i.e. it neither reflects nor transmits any of the incident radiation; consequently, if it is heated, it must emit radiations of all frequencies. No actual black surface fulfils these requirements entirely, so that a 'black body' to satisfy them must be produced artificially. If a sphere has a small hole in its side, and energy, in the form of radiation, enters that hole, the chances of it ever escaping again are very remote on account of the numerous reflexions taking place inside the sphere: at each reflexion a certain fraction of the energy is absorbed, so that eventually only a negligible quantity remains. If such a body is made hot [the material of the walls must be capable of withstanding the high temperatures to which they may be subjected] radiations of all wave-lengths will proceed from the aperture, i.e. such a body becomes an ideal radiator, and the radiation from it is known as ***'black body' radiation.*** ANDRADE, in his book, *The Mechanism of Nature*, gives the following illustration. 'For instance, an open window in a white house-front looks a perfectly black square on a sunshiny day: the sunshine is reflected from the white wall, which looks bright, but, passing through the hole into the room, is weakened at every encounter with objects there, and very little escapes again out of the window. The glowing heart of a furnace is an ideal radiator, for it is practically a small hole surrounded by glowing bodies all at one high temperature.'

'The paradox of the term "black body" appears when we consider what happens when we heat the walls of our iron vessel red hot, or even white hot. A bright light comes out of the hole, and yet we call this "black body radiation." All that is meant is that it is the kind of radiation which comes from a body that, since it absorbs all radiation that falls on it, presumably sends out, when heated, as much radiation of every kind as possible. The term "complete radiation" or "full radiation" probably expresses to the layman more clearly what is meant, but the term "black body radiation" is so widely used—and gives rise to so much misunderstanding—that this word of explanation has been offered.'

Distribution of Energy in the Spectrum of a Black Body.—In 1800, HERSCHEL, in examining experimentally the distribution of energy in the solar spectrum, discovered the existence of the invisible infra-red rays. He detected them by their heating effect on a thermometer placed beyond the red end of the sun's spectrum. He also discovered that the maximum calorific effect was situated in the infra-red region. Earlier investigators had located this position in the red (crown glass prism) and yellow (water). These differences are attributable to the absorption of energy in the material of the prism. MELLONI, using a prism of rock salt, found the maximum energy in the infra-red. In addition to the effect produced by selective absorption in the prism, it must be remembered that the energy distribution will depend on the dispersion produced. To avoid this difficulty a normal spectrum should be employed—this is a spectrum in which the deviation is directly proportional to the wave-

length. This cannot be done with a prism, so that the results of the energy distribution in a spectrum must always be corrected for this effect, i.e. the distribution of energy in a normal spectrum is calculated from that found experimentally in a spectrum which is not normal.

LUMMER and PRINGSHEIM, amongst others, investigated the distribution of energy in the spectrum of a black body. The radiation was obtained from a uniformly heated cylinder in which there was a small aperture, the temperature being measured by means of a thermocouple. The radiation was focused on a slit by means of silvered concave mirrors, and then fell upon a fluorite or quartz prism. The use of lenses was debarred on account of selective absorption in them. The energy between two neighbouring wave-lengths was measured by means of a vacuum bolometer. The first bolometer was constructed by LANGLEY in 1881. The working part of this instrument consists of a strip of thin platinum resembling a grating. It is covered with lamp-black. The grating is then placed in one arm of a Wheatstone bridge. A similar grating, but protected from all radiation, forms part of another arm of the bridge—it is termed the compensating resistance. A balance is obtained by varying a resistance in series with this compensating resistance, the two other arms being equal [compare the Callendar-Griffiths bridge]. When heat falls on the active part of the bolometer the balance is destroyed and the current through the galvanometer is a measure of the intensity of the radiation incident upon the bolometer.

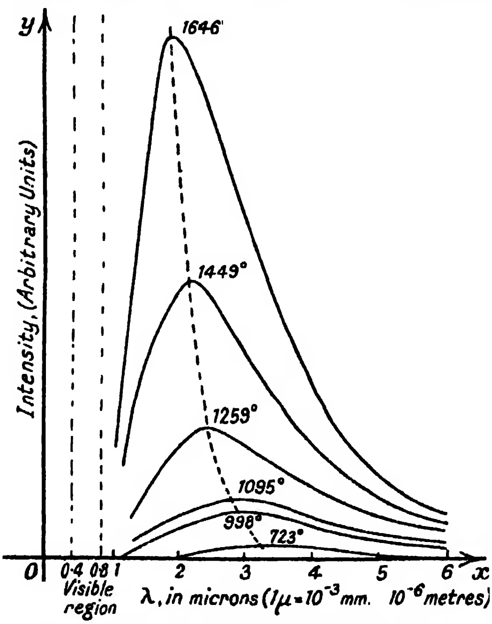


FIG. 16·9.—Distribution of Energy in the Spectrum of a Black Body.

Since the strip of the bolometer has a finite width it measures the energy due to radiations over a small range of wave-lengths.

The results are exhibited in Fig. 16·9. The ordinates are the intensities and the abscissæ the wave-lengths in microns (μ), [1 $\mu = 10^{-3}$ mm.]. The total energy E, emitted per second, for a given temperature is

$$E = \int_0^\infty E_\lambda . d\lambda,$$

and this is represented by the area between the corresponding curve and the x-axis. This area is directly proportional to the fourth power

of the temperature of the body measured on the absolute or Kelvin scale of temperature—Stefan's law.

Wien's Displacement Law.—The curves shown in the above diagram indicate that as the temperature is raised the maximum heating effect moves towards the region of shorter wave-length. WIEN was able to show theoretically that

$$\lambda_m.T = \text{constant} = 0{\cdot}294 \text{ cm. deg. K},$$

where λ_m is the wave-length corresponding to the maximum value of E. This is known as Wien's displacement law.

The Stefan-Boltzmann Law.—We have already seen how Stefan, in 1879, basing his argument on an experiment due to TYNDALL, suggested that the total energy emitted per second from a hot body was directly proportional to the fourth power of its temperature on the Kelvin scale. To express the Boltzmann-Stefan law mathematically we have to consider unit area (generally 1 cm.^{2}) of the hot body. Then if E is the energy (ergs) emitted per second from that unit area (1 cm.^{2}), the above law states that

$$E = \sigma T^4,$$

where σ is known as Stefan's constant. The mean value of σ obtained experimentally is

$$(5{\cdot}74 \pm 0{\cdot}02) \times 10^{-5} \text{ erg. cm.}^{-2} \text{ sec.}^{-1} \text{ deg.}^{-4} \text{ K.}$$
$$= (5{\cdot}74 \pm 0{\cdot}02) \times 10^{-12} \text{ watt. cm.}^{-2} \text{ deg.}^{-4} \text{ K.}$$

If we consider the energy due to radiations whose wave-lengths lie between λ and $\lambda + \delta\lambda$ to be $E\lambda.\delta\lambda$ per second per unit area, then the total energy emitted per unit area per second is given by $E = \int_0^\infty E_\lambda . d\lambda$.

This, by Stefan's law, is σT^4, where σ is the Stefan-Boltzmann constant. It is associated with the name of Boltzmann since he established Stefan's law theoretically, the clue to his argument having been provided by Maxwell, who showed that all radiations exert a pressure on any surface upon which they are incident.

LUMMER and PRINGSHEIM, in 1897, verified the validity of this law for a black body over a large range of temperatures, but their apparatus is too complicated to be considered here.

Example.—Obtain an expression for the rate of fall of temperature of a black body of area A cm.^{2}, and mass m, when it is at a temperature $\theta°$ C., the temperature of the walls of the exhausted enclosure being $\theta_0°$ C.

Let s be the specific heat of the material. Then the heat lost from unit area per second is

$$-\frac{ms\dfrac{d\theta}{dt}}{A} \text{ cal.}$$

This must be equivalent to

$$\sigma[(\theta + 273)^4 - (\theta_0 + 273)^4] \text{ erg.}$$

$$\therefore \text{ Rate of cooling} = -\frac{d\theta}{dt}$$
$$= \sigma.\frac{A}{Jms}[(\theta + 273)^4 - (\theta_0 + 273)^4] \text{deg. sec.}^{-1}$$

Solar Radiation and the Solar Constant.—An interesting problem arising in connexion with solar physics concerns the rate at which the sun emits energy. The amount of such energy expressed in calories falling per minute on an area of 1 cm.^{2} placed normal to the rays and situated outside the earth's atmosphere is termed the ***solar constant.***

LANGLEY made the first reliable determination of this constant. In his work he made corrections for the selective absorption, i.e. the absorption of different rays to different extents, of the atmosphere. A diffraction grating was used to separate out the different wave-lengths and the heating effects of consecutive small parts of this spectrum were measured with the aid of a bolometer. He examined in this way the distribution of energy in the solar spectrum, (i) at noon and (ii) when the sun's rays passed through twice the thickness of air. The curves he obtained are shown in Fig. 16·10. Curve (iii) is constructed from (i) and (ii) by drawing ordinates such that, for example,

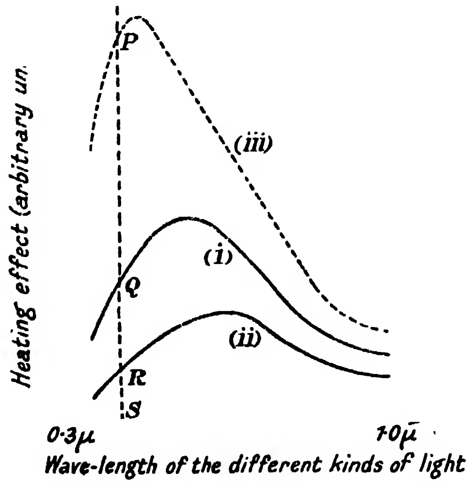


FIG. 16·10.—Langley's Curves for the distribution of energy in the Solar Spectrum.

$$\frac{PS}{QS} = \frac{QS}{RS}.$$

It may be shown that curve (iii) is the curve he would have obtained if the observations had been made outside the atmosphere.

Now the area under one of these curves is a measure of the total heat received. Langley found that

$$\frac{\text{Area of (i)}}{\text{Area of (ii)}} \quad 1{\cdot}57$$

so that the area of (iii) is 1·57 times that of (i).

Langley then measured the total heat received per minute per 1 cm.^{2} area of a surface normal to the sun's rays at noon on a clear day by using a special form of calorimeter known as an actinometer. When this result was multiplied by 1·57, he obtained 2·84 cal. cm.$^{-2}$ min.$^{-1}$ as the value of the solar constant.

Emissive Power.—The rate at which heat is lost from the surface of a body depends, as we have seen, upon the nature of the surface, the difference between its temperature and that of its surroundings, and on the material which constitutes the given body.

The emissive power of a surface is defined as the ratio of the amount of radiation emitted per unit time by unit area of the surface to the amount emitted per unit time by unit area of a perfectly black body, the emissions taking place under identical conditions. The emissive powers of different surfaces may be compared by the following method due originally to PROVOSTAYE and DESAINS :—A thermopile, T, Fig. 16·11, connected to a galvanometer, G, is placed about 50 cm. from a Leslie's cube, L, containing boiling water, the vertical sides of which are coated with various substances whose emissive powers are to be compared. Two screens, M and M_1, having openings at their centres, are placed as shown. The outer surfaces of these screens are covered with lamp-black while the inner ones are polished. By this arrangement any radiation from L upon the outer surface of M is absorbed while any radiation from M to T is diminished. M_1 also diminishes this, and at the same time prevents any radiation which may fall on its outer surface from extraneous sources from being reflected towards T. The currents in the galvanometer are proportional to the emissive powers of the surfaces responsible for the radiation.

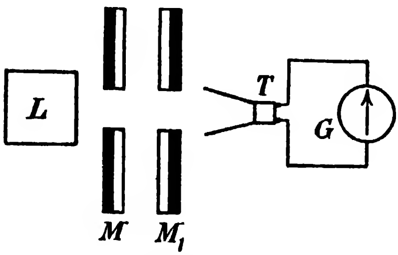


FIG. 16·11.—Comparison of Emissive Powers.

Absorption of Radiation.—Let a quantity of energy equal to Q fall on a surface every second and suppose that a quantity Q_1 is absorbed in the same time. The ratio $\frac{Q_1}{Q}$ is termed the ***coefficient of absorption*** or ***absorptive power*** of the surface. Relative values of the absorptive powers of different surfaces may be compared by a method first adopted by PROVOSTAYE and DESAINS, but a determination of the absolute value of the coefficient of absorption is difficult. A thermometer having its bulb coated with the substance under investigation is placed inside an enclosed box and a convex lens is employed to cause radiation to fall on the bulb. The bulb eventually assumes a steady condition in which the heat gained by absorption is equal to that lost by radiation: call this temperature θ_1. The thermometer is now warmed to a somewhat higher temperature than θ_1 and the rate at which it cools is observed and a cooling curve constructed. From this curve the rate of cooling at the temperature θ_1 is deduced. Let this rate of cooling be α_1. Then the heat lost per second by the

bulb is $m\alpha_1$ cal., where m is the thermal capacity of the bulb. This may be written $Jm\alpha_1$ erg.sec.$^{-1}$, where J is the mechanical equivalent of heat. Under the steady conditions here obtained, this is equal to the heat absorbed per second by the bulb, viz. A_1Q where A_1 is the absorptive power of the substance on the bulb, and Q is the amount of radiation [erg.] incident upon it per second, i.e.

$$Jm\alpha_1 = A_1Q.$$

Similarly, when a second substance is on the bulb,

$$Jm\alpha_2 = A_2Q,$$

hence

$$\frac{A_1}{A_2} = \frac{\alpha_1}{\alpha_2}.$$

To Verify directly that the Emissive Power of a Surface equals its Absorptive Power.—For this purpose the apparatus shown in Fig. 16·12 may be used. It is a modern form of RITCHIE'S apparatus and consists of a Leslie's cube filled with water which is kept boiling by passing a large current through the heating coil A. This coil

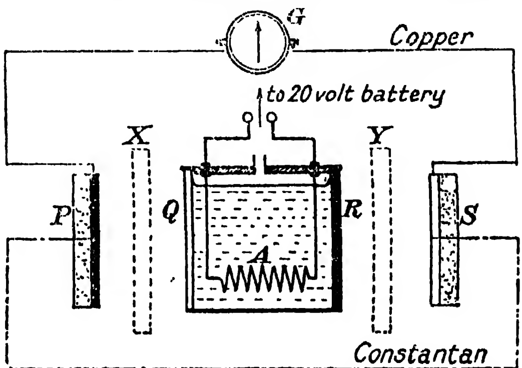


FIG. 16·12.—Modern Form of Ritchie's Apparatus.

should preferably be wound on a mica frame and the leads to it pass through ebonite blocks in the lid of the cube. The two surfaces Q and R of this cube (made of copper) are polished and lamp-blacked respectively. P and S are two thin copper sheets of the same size and thickness and arranged at equal distances from Q and R respectively. P is lamp-blacked, while S is polished. Immediately behind P and S are sheets of asbestos to assist the retention of any heat received by these plates. At the centres of P and S are soldered the ends of a constantan wire. A galvanometer G is connected to copper wires leading from the edges of the plates P and S. The galvanometer is then in series with a copper-constantan thermocouple. X and Y are two wooden screens which are removed when the water boils steadily. When this is done the galvanometer remains undeflected showing that there is no temperature difference between the junctions of the thermocouple, i.e. the heat received per second by S and P is the same.

A more instructive method of carrying out this experiment is as follows: one of the screens, say X, is removed and the galvanometer deflexion observed. When the screen Y is subsequently removed, the above deflexion is reduced to zero, showing that the amounts of heat received by P and S are equal.

Let H be the heat (ergs) emitted from R per second. Then the heat received per second by S is αAH, where A is its absorptive power and α a coefficient depending on the disposition of R and S. Let E be emissive power of Q (and S). Then the heat emitted from Q per second is EH—cf. the definition of emissive power on p. 340. Since the disposition of P and Q is the same as that of R and S the amount of heat received from Q by P in one second is $\alpha EH \times 1$, since the absorptive power of a lamp-blacked surface is unity. Hence, since the heat received by S is equal to that received by P in the same time, we have $\alpha AH = \alpha EH$, i.e. $A = E$.

Surface Emissivity.—The ***surface emissivity*** of a body is defined as ***the quantity of heat lost per unit time per unit area of its surface per degree excess temperature.*** [N.B.—The heat is lost by radiation, by conduction, and by convection.] To determine the surface emissivity of copper the apparatus shown in Fig. 15·17, of p. 318, may be used. The rate at which the temperature of the body is changing at any instant is ascertained in the manner previously indicated. Let this be α degrees per second when the temperature excess is θ. The heat lost per second under the above condition is $(MS + ms)\alpha$. This is equal to $4\pi r^2\sigma\theta$, where σ is the surface emissivity of copper. Hence σ is given by

$$\sigma = \frac{(MS + ms)\alpha}{4\pi r^2\theta} \text{ cal. sec.}^{-1} \text{ cm.}^{-2} \text{ deg.}^{-1} \text{ C.},$$

if, as usual, the temperature is expressed on the centigrade scale, and the other quantities in c.g.s. units. When the excess temperature is large, σ is not a constant for a given surface.

The Dewar Flask.—The Dewar or thermos flask was designed for the specific purpose of diminishing the rate of exchange of heat between the contents of the flask and its surroundings. Originally it was designed for storing liquefied gases. The vessel, Fig. 16·13, is generally made of glass, the space between its double walls being exhausted to a very high vacuum. This constitutes the best-known obstacle to the transference of heat by conduction and convection; it is only radiant energy which can pass from one wall to the other across the vacuum, and the rate at which this occurs is diminished by coating the interior walls with metallic silver. The surface of the silver in contact with the glass assumes a high degree of polish, so that the outer wall of the inner vessel becomes a poor radiator of energy, whilst the inner wall of the

outer vessel reflects any radiant energy it does receive. Silvering the walls only, however, would not make these flasks efficient, for the air remaining between the walls would still assist in the transfer of heat between the contents of the flask and the surroundings. Similarly, if the walls are not silvered but the space between them is evacuated, the efficiency of the flask is low, since the rate of transfer of heat by radiation has not been reduced. The mouth of the flask is made narrow, and is usually closed with a cork [but not when liquefied gases are inside the flask]. This cork further reduces the amount of heat reaching the interior of the flask through conduction or convection. To protect the flask from mechanical shocks it is usually supported in a metal case by means of a spring and suitable soft non-conducting material.

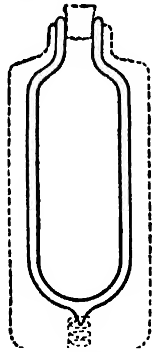


FIG. 16·13.—A Dewar Flask.

EXAMPLES XVI

1.—State Newton's law of cooling, and discuss its validity. Describe how you would proceed to compare the emissivities of the surfaces of a blackened and a polished copper calorimeter of the same dimensions.

2.—Explain the term diathermancy. How would you measure it for a plate of glass ? Discuss the effects of varying (1) the thickness of the plate, (2) the temperature of the source, on the diathermancy of the material. (L.)

3.—A metal sphere of thermal capacity 6·5 cal. deg.$^{-1}$ C. is observed to be cooling at a rate of 0·5° C. per minute when its temperature is 50° C. above that of its surroundings. If the sphere is 3 cm. in diameter, calculate the thermal emissivity of the surface of the sphere.

4.—A copper sphere, whose thermal capacity is 4 cal. deg.$^{-1}$ C., receives heat energy at a steady rate and finally attains a temperature of 60° C. The temperature of the sphere is then raised, the supply of energy removed, and, from a cooling curve constructed in the usual manner, it is found that at 60° C. the sphere cools at a rate of 5 degrees in 2 minutes. Obtain a value for the rate at which heat energy was supplied to the body in the first part of the experiment.

5.—State *Stefan's law of radiation.*

Calculate the maximum net rate of loss of heat by radiation from a sphere of 10 cm. radius at a temperature of 200° C. when the surroundings are at 20° C., if Stefan's constant is $5{\cdot}7 \times 10^{-12}$ watt. cm.$^{-2}$ deg.$^{-4}$ K.

6.—Define *surface emissivity.*

Assuming Newton's law of cooling, compare the rates of loss of heat in a vacuum of two copper spheres whose radii are 8 cm. and 4 cm., their temperatures being respectively 15° C. and 7·5° C. above that of the enclosure in which they are placed. Compare also the

rates of fall in temperature of the two spheres. [Assume that both spheres conduct heat perfectly and that their surfaces are identical in nature.]

7.—Explain how you would demonstrate that the heating effect of the sun's radiation is not confined entirely to its visible spectrum. Explain the principle of the detecting instrument you would use in this demonstration.

8.—State *Stefan's law of radiation* and give a general account of the distribution of energy in the spectrum of a black body.

A solid copper sphere cools at the rate of 1·4° C. min.$^{-1}$ when its temperature is 127° C. Obtain a value for the rate at which a solid copper sphere of one-half the radius will cool when its temperature is 227° C. if, in both instances, the surroundings are maintained at 27° C.

9.—A block of copper, of mass 100 gm. and specific heat 0·10 cal. gm.$^{-1}$ deg.$^{-1}$ C., which can be heated by electrical energy supplied at a rate of 2·09 watt., is suspended in an evacuated enclosure, the walls of which are maintained always at the same temperature as the copper block. Explain how these particulars enable one to ascertain the manner in which the temperature of the block varies when the electrical energy is switched on to it.

10.—Describe how you would investigate the relative importance of the various methods by which a small calorimeter containing hot water loses heat, and state the general nature of the results you would expect to obtain from your experiments.

11.—A vessel of total thermal capacity 1000 cal. deg.$^{-1}$ C. has a blackened surface of area 500 cm.^{2}. It is surrounded by an enclosure, the blackened walls of which are maintained at 27° C. Find a value for the rate of fall in temperature of the vessel and its contents when their common temperature is 87° C. Stefan's constant may be taken as $5{\cdot}7 \times 10^{-5}$ erg.cm.$^{-2}$ sec.$^{-1}$ deg.$^{-4}$ and it may be assumed that the contents of the vessel have a very high thermal conductivity.

12.—'*The heat lost by radiation from a full radiator is proportional to the difference of the fourth powers of the absolute temperatures of the body and the enclosure.*' Justify this statement and use it to find an expression for the rate of fall in temperature of a body of mass m, surface area A, and whose material has a specific heat s, in terms of ϕ, the excess of temperature of the body over T_0, the temperature of the enclosure.

13.—A copper sphere is heated and then allowed to cool while suspended in an enclosure whose walls are maintained at a constant temperature. When the temperature of the sphere is 86° C. it is cooling at a rate of 3 deg. C. min.$^{-1}$; at 75° C. the rate of cooling is 2·5 deg. C. min.$^{-1}$. Calculate values for (*a*) the temperature of the walls of the enclosure, and (*b*) the temperature of the sphere when it is cooling at the rate of 1 deg. C. min.$^{-1}$.

14.—A copper calorimeter of mass 265 gm. contains 450 cm.^{3} of water and cools through 6° C. in 5 minutes. How long would it take the same calorimeter to cool through the same range of temperature and under the same conditions if it contained 450 cm.^{3} of a liquid of density 0·8 gm.cm.$^{-3}$ and specific heat 0·75 cal. gm.$^{-1}$ deg.$^{-1}$ C.? The specific heat of copper may be taken as 0·1 cal. gm.$^{-1}$ deg.$^{-1}$C.

PART III

OPTICS

CHAPTER XVII

GENERAL INTRODUCTION

The earliest investigations concerning the nature and behaviour of light were probably made by the ancient Egyptians. Their work was followed by that of the Greeks; the early physicists [PYTHAGORAS and his followers, 580 B.C.] believed that the eye simulated an octopus. The tentacles, which were supposed to project from the eye, seized an object and illuminated it. DEMOCRITUS (510 B.C.) held the opposite view, for according to him the images produced on the retina of the eye arose from something which was emitted from the object. PLATO [430 B.C.] tried to combine the two theories; he regarded light as a phenomenon produced by the collision of emanations from both the eye and the object. It is believed that Plato and his disciples enunciated two of the fundamental laws of light, viz. that light travels in straight lines, and that when it is reflected from a mirror the angle of incidence is equal to the angle of reflexion. ARCHIMEDES—'The Father of Physics'—who lived about 287 B.C. was a capable experimentalist, and when the Romans attacked Syracuse in 212 B.C. it is said that this ancient philosopher constructed huge mirrors with which he set fire to the enemy ships which were lying at anchor. By A.D. 100 PTOLEMY had become acquainted with the bending of light which occurs when light passes over the boundary between two transparent media. From then onward the progress of this science was slow, but it is interesting to learn that our own countryman, ROGER BACON [1214], was interested in optics, and that his knowledge of burning glasses [lenses] and mirrors was clear. Then came COPERNICUS [1473], GALILEI GALILEO [1564], and KEPLER [1571], to whom the nature of light began to reveal itself. Galileo invented the telescope and made many contributions to the science of optics.

SIR ISAAC NEWTON [1642–1727] carried out many researches regarding optics and before long showed that white light was hetero-

geneous. He regarded a beam of light as a train of corpuscles which impinged upon the retina and stimulated the sensation called vision. Newton had shown that all material bodies attracted one another, so he naturally supposed that these light corpuscles were attracted by a transparent medium—this attraction was the cause of refraction. In order to account for the reflexion of light Newton developed his so-called 'Theory of Fits,' according to which some of the corpuscles were attracted by the medium and some repelled.

At the beginning of the nineteenth century YOUNG and FRESNEL introduced the wave theory of light, confirming their theory by actual experiment. They showed that light could bend round corners and that this could be accounted for if light consisted of waves. Since then the wave theory has been developed in the hands of MAXWELL, KELVIN, and others. Whether or not the wave theory is to be the ultimate truth regarding the nature of light is not known; at present there are several ideas extant, but they cannot be discussed in this book.

The Æther and Light Waves.—In the Wave Theory of Light the object, which is seen, is the source of the light waves, and some medium is supposed to be necessary for the propagation of these waves. It is at once obvious that the air is not the transmitting agency for the stars are visible although there is every reason to believe that the interstellar space is void of matter. Young imagined that an all-pervading medium, the *æther,* was responsible for the conveyance of luminous energy. Light waves are similar to those which spread over the surface of a pond into which a stone has been thrown; small objects floating on the water merely move up and down while the waves pass by—the objects are not carried forward although the waves trável in that direction.

When the light waves are incident upon any small object the light is scattered—a beam of sunlight entering through some small hole into a darkened room is not visible except for the small motes present in its path. These dust particles become visible because they scatter the light which is incident upon them, thereby indicating the path of the beam of light. When the beam from a searchlight cannot be seen, if an object, such as an aeroplane or distant ship, comes into the beam that object is vividly illuminated. Such phenomena show that a beam of light is not visible unless it is incident upon some object.

Rays and Pencils of Light.—The path along which light energy travels is called a *ray.* Since light consists of very small waves, rays have no real physical existence, but the conception of a ray is useful in that it simplifies our calculations. The branch of this subject which deals with rays is called *Geometrical Optics* to distinguish it from *Physical Optics,* in which the wave motion

is considered. It must be carefully noted, however, that a result which has been obtained by means of geometrical optics is not necessarily true; unless the same result can be inferred from physical optics the result must be viewed with suspicion. When a bundle of rays proceeds from the source in some particular direction that bundle is generally referred to as a ***beam*** or ***pencil of light.*** If the light rays tend to open out as they proceed from the source, the beam is said to be ***divergent***; if the rays tend to pass through a point then the beam is ***convergent***; if the rays remain parallel the beam is termed a ***parallel*** one.

The Rectilinear Propagation of Light.—That light rays travel in straight lines in a homogeneous medium is the foundation upon which the science of geometrical optics has been built. To show that light travels in straight lines the following experiment may be made. If three screens (metal sheets) are each pierced with a small hole, and held between a source of light and the eye, the source is only visible if the holes are collinear, i.e. in the same straight line. A slight displacement of any one of the screens and the source is no longer visible. Later on we shall learn that light waves 'do bend round corners' but that it is because their wave-lengths are so short and the amount of bending therefore small, that such effects were not noticed until about the end of the eighteenth century.

Shadows.—The formation of a shadow is a natural consequence of the fact that light travels in straight lines. If a pointolite lamp is placed at some distance away from a vertical brass tube several inches in diameter, a well-defined shadow is found on a screen placed a little distance away from the tube. If several such point sources of light are used, then each one casts a distinct shadow. When the pointolite lamp is replaced by an ordinary metal filament lamp *a b*,

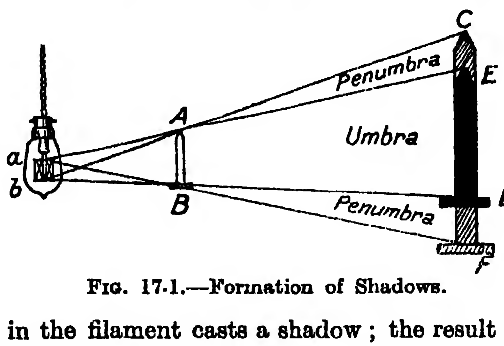


FIG. 17·1.—Formation of Shadows.

Fig. 17·1, each point in the filament casts a shadow; the result is that many shadows of the object are produced and these may, or may not, overlap. The dark region from which all direct light is excluded is called the ***umbra,*** ED; where the ***half-shadow*** or ***penumbra*** region occurs the screen is receiving light from some fraction of the source and is therefore partly illuminated, CE, FD. Beyond the region of the penumbra the screen is fully illuminated.

Eclipses.—Solar and lunar eclipses are the results of the formation of shadows by the moon or earth, the sun being the source of light. If the moon, during its journey in space, moves into a position between the sun and the earth, a portion of the sunlight falling upon the earth is intercepted and there is an eclipse of the sun. The

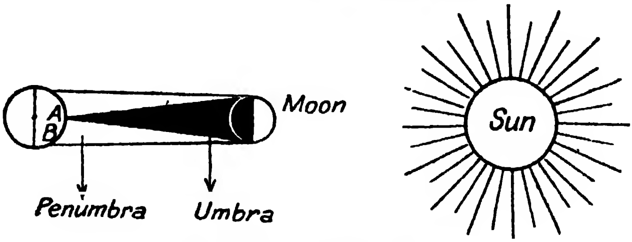


FIG. 17·2.—A Solar Eclipse.

state of things during a solar eclipse is shown in Fig. 17.2. If an observer is in the umbra, A, the eclipse is total—at regions in the penumbra, B, the phenomenon of a partial eclipse may be seen.

An eclipse of the moon occurs when the earth is in such a position that it intercepts the sunlight which would otherwise have rendered the moon luminous.

PHOTOMETRY

Light as a Measurable Quantity.—We have already seen that when light falls on the blackened surface of a thermopile a rise in temperature of this surface takes place. This rise in temperature is a measure of the energy in the light waves. Hence the total amounts of energy emitted by two light sources might be compared by placing them in turn at the same distance from a thermopile, and observing the deflexions of a galvanometer placed in series with the thermopile. The ratio of these deflexions is the ratio of the total energies emitted by the two sources in a given time, since the rise in temperature is proportional to the deflexion. Unfortunately, however, this ratio is not a measure of the comparative brightness of the two sources, for the brightness depends on the wave-length. Hence the calorimetric or physical method of comparing light sources must be replaced by a photometric or physiological test when the relative brightness of light sources is being estimated. Moreover, persons differ in their opinions regarding the brightness of various lights so that instruments must be used if the lights are to be compared accurately. An instrument for this purpose is termed a ***photometer,*** while this particular branch of optics is termed ***photometry.***

Light Standards.—Since, in the science of photometry, we have to compare the intensities of different light sources it becomes

imperative to select some standard source of light as a unit in which all other intensities may be expressed. This standard must satisfy the demands made upon all standards, viz. it should be constant, or at most only subject to slight variations which can be allowed for when the standard is in use, and it should be independent of the observer who sets up the standard providing he pays attention to certain specified details, i.e. it should be reproducible from specification. Moreover, the spectral distribution of its light [cf. p. 447] should approximate to that of the source compared with it.

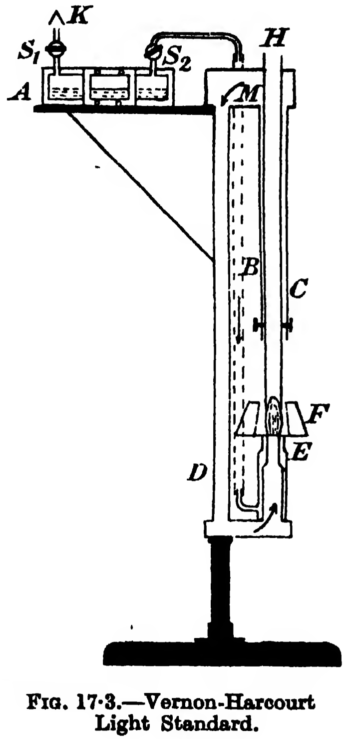


FIG. 17·3.—Vernon-Harcourt Light Standard.

The oldest form of standard is the spermaceti candle $\frac{7}{8}$ inch in diameter, having a mass of $\frac{1}{6}$ lb., and burning at a rate of 120 grains per hour. Variations in the shape of the wick and the fact that the luminosity of the flame is influenced by the water content and temperature of the air prevent this candle from fulfilling the requirements of a primary standard, so that it has been replaced by the VERNON-HARCOURT pentane standard. This is shown in Fig. 17·3. Liquid pentane, a highly inflammable substance, is contained in the 'saturator' A. A mixture of air and pentane vapour passes from A to the steatite burner E via the rubber tube B. Here it is burnt and the products of combustion escape up the tube H. These warm the air in the tube C surrounding H and this air passes into M where it is cooled; it then descends at a constant rate along D to the burner. The percentage of air in the pentane-air mixture is controlled by adjusting the stop-cocks S_1 and S_2 and the cone K. A regular evaporation of the pentane is established by heat passing along the bracket supporting the saturator. The flame is protected from draughts by a collar F and the top of the flame is hidden from view by the lower end of the tube H.

The special features of such a lamp are that the vapour which is burnt has a definite chemical composition and that the combustion

takes place under definite conditions. The flame is adjusted to a fixed height and an opening of definite size in a metal cone surrounding the flame permits only the light energy from a definite area of the flame to be used.

This lamp has an intensity equal to that of about ten spermaceti candles so that the unit of candle-power, in use until 1940, was defined as one-tenth that of a pentane standard lamp.

The standard of candle-power—*the international candle*—now used in photometric laboratories is one-sixtieth of the candle-power per square centimetre of a black body at the temperature of solidification of platinum ($1755 \pm 6)°$ C. The walls of the black body consist of fused thorium oxide.

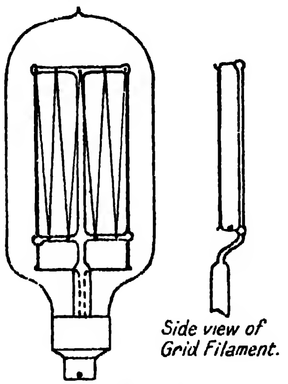


FIG. 17·4.—Light Standard N.P.L. Sub-standard Lamp.

The difference between the old and new units is small, but significant in accurate work.

For many purposes it is sufficient to use a sub-standard lamp such as that indicated in Fig. 17·4. This lamp has its filament in one plane and this is arranged perpendicular to the axis of the photometer bench when in use. When a definite constant voltage is applied to the lamp its intensity is very constant.

Definitions used in Photometry.—(i) ***Luminous Flux.*** Luminous flux, F, is defined as the rate of passage of light energy evaluated with reference to the luminous sensation it produces. The unit of luminous flux is the ***lumen***; this is the amount of luminous energy received per second by unit area of a sphere of unit radius when a uniform point source of one international candle is placed at the centre of the sphere.

(ii) ***Illumination.*** The illumination, E, at a surface is defined as the quantity of light falling on unit area of that surface per second.

If F is the luminous energy emitted per second by a uniform point source, then the illumination at a point on a surface normal to the direction in which the light is propagated and at a distance r from the source is measured by $F/4\pi r^2$, since this amount of energy is received per second by unit area of a sphere drawn round the luminous point as centre. The illumination at a surface is usually expressed in ***foot-candles*** or in ***metre-candles***. Thus, the statement that the illumination at a surface is four metre-candles implies that it is the same as if it were illuminated by four standard

candles placed at a distance of one metre from it. A metre-candle is termed a ***lux***

(iii) ***Candle-Power or Luminous Intensity.***—Let us assume that a sphere has been drawn round a luminous point emitting light equally in all directions, the luminous point being at the centre of the sphere. Then the amount of light falling on unit area of this sphere per second is directly proportional to the quantity of light emitted by the source in the same time. Now the ***candle-power*** or ***luminous intensity,*** T, of a source is numerically equal to ***the ratio of the quantity of light falling per second on unit area of such a sphere to the amount which falls from a point source of one international candle in the same position on the same area in the same time.*** The luminous intensity of a light source is expressed in candle-power.

The Inverse Square Law.—The fact that a light becomes fainter the more remote it is from an observer, is well known. The manner in which the light becomes fainter may be calculated as follows. Let S, Fig. 17·5, be a point source of light, or radiant, situated at the centre of a sphere A. Let F be the total quantity of light emitted per second by the source. The quantity of light falling per second on unit area of A, a sphere of radius r_a, is $\frac{F}{4\pi r_a^2}$. Similarly, when the first sphere is removed, the quantity of light falling on unit area of a sphere of radius r_b per second is $\frac{F}{4\pi r_b^2}$ Hence

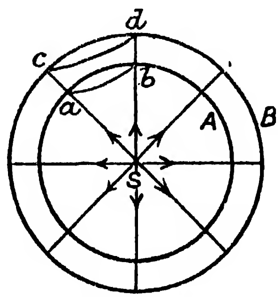


FIG. 17·5.

$$\frac{\text{Illumination at a point on the surface of A}}{\text{Illumination at a point on the surface of B}} = \frac{\frac{F}{4\pi r_a^2}}{\frac{F}{4\pi r_b^2}} = \frac{r_b^2}{r_a^2},$$

i.e. $\frac{E_a}{E_b} = \frac{r_b^2}{r_a}$ or, in general, $Er^2 = \text{constant}$,

i.e. the illumination of a surface varies inversely as the square of its distance from the source of light [strictly, only if the distance is large compared with the linear dimensions of the source which, in the above argument, has been assumed to be a point source].

The above formula may also be obtained from the fact that the quantity of light falling on a portion *ab* of the sphere A

per second would, in the absence of A, fall on an area *cd* of the sphere B, and that these areas, being geometrically similar, are proportional to the squares of the radii r_a and r_b respectively.

Illumination due to a Point Source.—If I is the candle-power of a point source of light placed at the centre of a sphere of radius r, the total amount of light energy falling normally per second on the surface of the sphere, area $4\pi r^2$, is $4\pi I$ lumens. The illumination, E, at any point on the surface of this sphere is therefore given by

$$E = \frac{4\pi I}{4\pi r^2} = \frac{I}{r^2}.$$

Hence, if two sources of candle-powers I_1 and I_2 are arranged so that they produce equal illuminations on screens at distances r_1 and r_2 from them respectively, then

$$\frac{I_1}{r_1^2} = \frac{I_2}{r_2^2}.$$

In comparing the illumination at two surfaces, it must be remembered that these illuminations are judged by the brightness of the surfaces, i.e. by the light reflected from them. It is therefore essential that the reflecting powers of the two surfaces should be equal, where, if Q_1 is the quantity of light reflected from a surface when a quantity Q falls upon it, the ratio $\frac{Q_1}{Q}$ is termed the ***reflecting power*** of the surface.

Brightness.—The illumination at a point on a surface receiving light energy does not depend on the nature of the surface. Thus a piece of white paper and a piece of black cloth lying side by side are equally illuminated since they receive per second per unit area the same amounts of light energy; but they appear very different to an observer. This difference is attributed to the fact that the above two surfaces behave differently with respect to the amount of incident light energy they 'reflect' to the eye. The ***brightness*** of a surface in a given direction is defined as the quotient of the luminous intensity of an element of the surface and the area of the element projected on a plane normal to the given direction. The ***unit of brightness*** is the candle. cm.$^{-2}$; this is often termed the ***lambert.***

The brightness of a surface is independent of the distance from which it is observed, i.e. the illumination of the retina is constant. This is because both the flux of light energy and the retinal image are inversely proportional to the square of the distance.

Rumford's Photometer.—This simple arrangement for comparing the candle-powers of two sources of light is shown in Fig. 17·6. In front of a white screen there is placed an opaque rod C. If M and N represent two sources of light to be compared, then two shadows are thrown on the screen placed behind the rod. It is desirable that the lines NA and MB should be equally inclined to the normal through C to the screen since the intensity of illumination depends upon the obliquity of the light rays [cf. p. 366]. The portion A of the screen, which would be the portion of the screen covered by the shadow of C when M is the only source of light present, is not completely dark when the two sources of light are present because it receives light from N; similarly, B receives light from M. The two sources of light are moved until the shadows A and B are just touching, and until the shadows cannot be differentiated from one another. The shadows are caused to touch

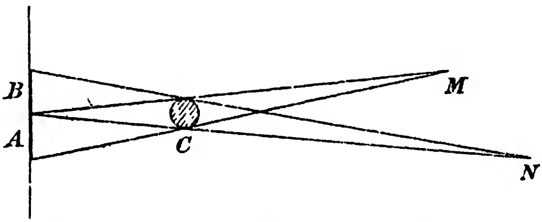


FIG. 17·6.—Rumford's Shadow Photometer.

since experience teaches that the equality may best be judged under these circumstances. When such conditions have been obtained

$$\frac{I_m}{BM^2} = \frac{I_n}{AN^2}, \text{ i.e. } \frac{I_m}{I_n} = \frac{BM^2}{AN^2},$$

where I_m and I_n are the candle-powers of the lamps at M and N respectively.

It is sometimes maintained that Rumford did not use a photometer of this very simple type. His first photometer, invented in 1794, did consist of one rod $\frac{1}{4}$ in. in diameter and 6 in. high placed in front of a white paper screen. The lamps were placed so that they could be moved along wooden grooves—these were equally inclined to the screen, i.e. as Rumford says, 'One light is precisely in the line of reflexion of the other.' This condition was 'easily performed by actually placing a piece of looking-glass, 6 or 8 in. square, flat upon the paper and observing by means of it the real lines of reflexion of the lights from that plane.' The glass was afterwards removed. The precaution taken by Rumford must always be observed, for the illumination depends upon the obliquity of the incident rays. The observer sat with the rod between himself and

the screen and moved the lamps with the aid of cords passing over pulleys.

Rumford found it very inconvenient to compare two shadows projected by the same cylinder as these were either necessarily too far from one another to be observed with certainty, or when they were nearer they were partly hidden from the eye by the cylinder. To remedy this inconvenience Rumford used two cylindrical rods R_1 and R_2, Fig. 17·7 (*a*), which were placed in front of a white paper screen, S. The rods and paper were placed inside a blackened box, the front of which alone was open. In this way extraneous light was excluded. Four shadows were then obtained—the two outer ones were effectively removed by narrowing the field of view. For this purpose the outer portions of the paper were coated with a dull black paint. The cylinders were made of brass and $2\frac{1}{5}$ in.

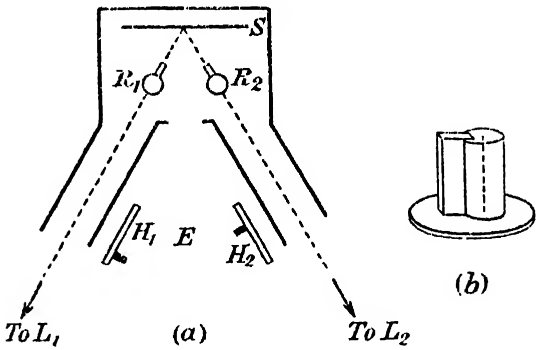


FIG. 17·7.—Rumford's Shadow Photometer—Original Pattern.

high: they were painted black. Each cylinder had a vertical flange or wing attached, $\frac{11}{20}$ in. wide and $\frac{1}{18}$ in. thick, cf. Fig. 17·7 (*b*), and could be rotated about a vertical axis. The wings commonly lay in the middle of the shadows of the cylinders, but when necessary one of the shadows could be increased in extent by rotating the cylinders about the vertical axes till the wing emerged from the shadows of the cylinders and their own shadows appeared. The rotation was continued until the shadows cast by the lamps L_1 and L_2 were just in contact along the centre of the screen. They were viewed from E. Equality of illumination was obtained by moving the lamps by means of cords operated by handles at H_1 and H_2.

Rumford says, 'If these shadows should be found to be of unequal densities, then that light whose corresponding shadow is densest, must be removed further off, or the other must be brought nearer, till the densities are equal.' Since the shadows are equally

dense, the intensities of illumination due to the lamps at the screen must be equal. The usual relationship between the candle-powers of the lamps and their distances from the screen (measured along the 'lines of reflexion' on which they were placed) is then applicable.

Bunsen's Grease-Spot Photometer.—The modern form of BUNSEN's photometer, Fig. 17·8, consists essentially of a grease spot on a piece of paper, the two sources of light which are to be compared being placed one on each side of it and on a common normal to the paper. The plane mirrors M_1 and M_2 are inclined to the grease spot [shown dotted] so that an observer may view both sides of the greased paper at once.

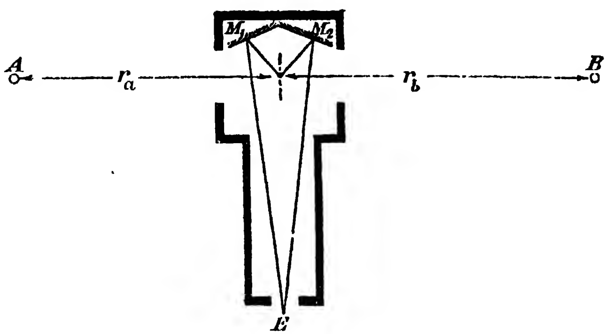


FIG. 17·8.—Bunsen's Photometer.

If I_a and I_b are the candle-powers of the sources being compared, then the illuminations at the screen due to the lamps are

$$\frac{I_a}{r_a^2} \quad \text{and} \quad r_b^2$$

respectively, i.e. these fractions give the amount of light falling per second on unit area of the screen. Let α_1 be a quantity $(0 < \alpha_1 < 1)$ such that α_1 is the fraction of the light falling on the opaque portion of the disc which is received by the observer; then α_1 will depend on the reflecting power of the surface, the obliquity of the screen with respect to the line of vision, and on the diameter of the pupil of the observer's eye. Then if the eye is placed in a symmetrical position with respect to each side of the disc, the amounts of light entering the eye, E, per second are

$$\alpha_1 I_a \quad \text{and} \quad \frac{\alpha_1 I_b}{r_b^2},$$

from unit areas of the two sides of the opaque portion.

Since the reflecting power of the opaque portion of the disc is different from that of the waxed portion it follows the light reflected from unit area of this per second and received at E will be

$$\frac{\alpha_2 I_a}{r_a^2} \text{ and } \frac{\alpha_2 I_b}{r_b^2},$$

where $\alpha_2 \neq \alpha_1$. But some light is also transmitted by the waxed portion. If unit quantity of light falls on a unit area of the waxed portion let β be the fraction of this which is received at E after transmission through the disc. Then, on the l.h.s., the amount of light received from unit area of the waxed portion of the disc per second is

$$\frac{\beta I_b}{r_b^2}.$$

Hence the total light per second from unit area of the l.h.s. of the waxed portion of the disc is

$$\frac{\alpha_2 I_a}{r_a^2} + \frac{\beta I_b}{r_b^2}.$$

The fraction
$$\frac{\dfrac{\alpha_2 I_a}{r_a^2} + \dfrac{\beta I_b}{r_b^2}}{\dfrac{\alpha_1 I_a}{r_a^2}}$$

may be considered as a measure of the contrast between the waxed and opaque portions of the disc on the l.h.s.

Similarly, on the r.h.s., the contrast may be expressed by the fraction

$$\frac{\dfrac{\alpha_2 I_b}{r_b^2} + \dfrac{\beta I_a}{r_a^2}}{\dfrac{\alpha_1 I_b}{r_b^2}}.$$

The experimental determination of the ratio of the candle-powers of two lamps with the aid of the Bunsen grease-spot photometer therefore consists in adjusting the positions of the lamps with respect to the screen, until there is equality of contrast between the waxed and opaque portions on each side of the disc. Then

$$\frac{\dfrac{\alpha_2 I_a}{r_a^2} + \dfrac{\beta I_b}{r_b^2}}{\dfrac{\alpha_1 I_a}{r_a^2}} = \frac{\dfrac{\alpha_2 I_b}{r_b^2} + \dfrac{\beta I_a}{r_a^2}}{\dfrac{\alpha_1 I_b}{r_b^2}},$$

i.e.
$$\frac{I_a}{r_a^2} = \frac{I_b}{r_b^2}.$$

To commence the above experiment, the lamps A and B are fixed about two metres apart and the grease spot moved until the above equality of contrast exists. The ratio of the candle-powers may then be computed.

[The remarks made above with reference to the light reflected from the opaque portion of the disc apply to any photometer in which such a surface is used, and will not be repeated.]

Abney's Variable Sector Photometer.—Instead of varying the relative distance of the two sources from the photometer disc, a matter of some inconvenience when one source is much stronger than the other, ABNEY kept them at equal distances from it, and reduced the effective intensity of the stronger light by means of a rotating sector situated in front of the light source. This sector consisted of a circular disc from which a sector had been removed. The angular width of this opening could be varied whilst the disc was actually running.

The experiment therefore consisted in adjusting the angular opening in the disc until, with the two sources at equal distances from the disc of a photometer, equality of illumination (or of contrast in the case of a grease spot photometer) was obtained. If n is the measure of the angular opening in radians and I the candle-power of the standard source behind the sector, then its effective candle-power is

$$\frac{n}{2\pi}\mathrm{I},$$

and this is the candle-power of the light which is being compared with the standard lamp.

To test the accuracy of the above factor $n/2\pi$, the candle-power of a lamp is compared with that of a standard lamp. Let these candle-powers be I_1 and I_s respectively, the distances of the lamps from the photometer disc being r_1 and s. Then

$$\frac{\mathrm{I}_1}{r_1^2} = \frac{\mathrm{I}_s}{s^2}.$$

The sector is then placed in front of one of the lamps—say S_1—and rotated rapidly: there is no question of any flicker [cf. p. 362]. The distance of S_1 from the disc is altered until the field of view is as before. Let r_2 be this distance. If I_2 is the effective candle-power of S_1 when this is behind the sector, then

$$\frac{\mathrm{I}_2}{r_2^2} = \frac{\mathrm{I}_s}{s^2}$$

i.e.

$$\frac{\mathrm{I}_2}{\mathrm{I}_1} = \frac{r_2^2}{r_1^2}.$$

The reduction factor for the disc is therefore r_2^2/r_1^2. This should be equal to $n/2\pi$.

The sector may therefore be standardized for various settings of its aperture and then used in connexion with a photometer. The above standardization is only necessary when n is small.

The Lummer-Brodhun Photometer.—This photometer is of great use in accurate work on photometry and it is depicted in Fig. 17·9 (a). A screen S is made of some pure white material, such as barium sulphate, and its opposite faces are illuminated by the two lamps situated at A and B, which are to be compared. Two mirrors are placed at M_1 and M_2, whilst P is a combination of two right-angled prisms. The outer portion of the base of one of

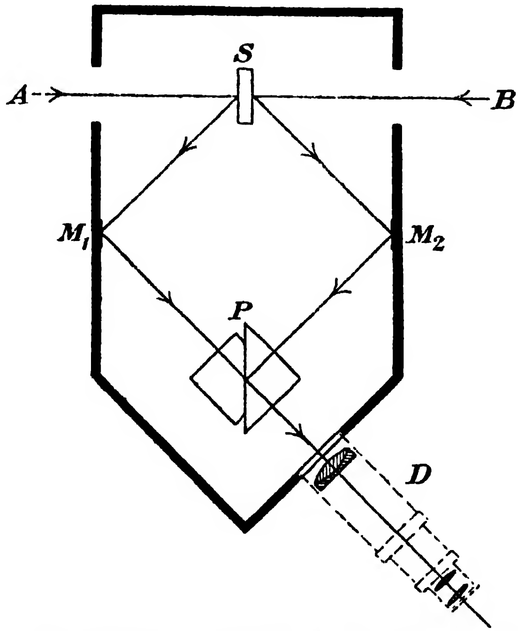


Fig. 17·9 (a).—Lummer-Brodhun Photometer.

these prisms has been ground away so that the prisms are only in contact over the central region of their bases. Good optical contact is here obtained by coating this region with Canada balsam. The complete outfit is enclosed in a blackened box. Light from A is scattered at S, some of it falling on to M_1 where it is reflected on to the nearer face of the prism, P. The light enters the prism and that portion which falls on the central region is transmitted through the prism, whilst the other portion is reflected from that part of the base of the first prism which is not in contact with the second one. Light from B follows a similar path, and the portion entering the low power (or long focus) microscope D is that which

does not traverse the central region—see Fig. 17·9 (*b*). There are now two light beams entering the microscope D, so that the field of view consists of two unequally illuminated portions when the microscope is focused upon the central portion of the bases. The positions of the lamps are adjusted so that the two sections of the field of view are equally bright. The usual relationship holds. The head attached to S is then rotated through 180° and the

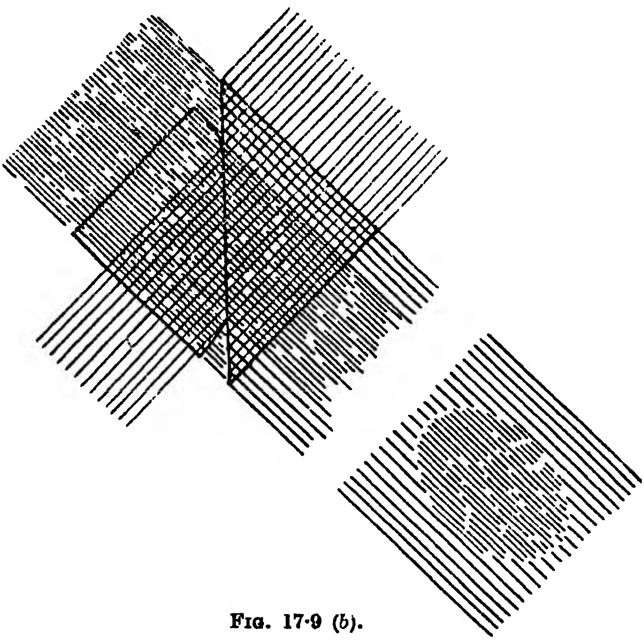


FIG. 17·9 (*b*).

experiment repeated. The mean of the results thus obtained eliminates any effects arising from the fact that the two sides of S may not be identical. This photometer is fundamentally more scientific than the others which have been described, principally so because the sources are both viewed with one eye, so that any difference which may exist between the eyes of an observer is of no consequence.

Heterochromatic Photometry, Flicker Photometers.—Heterochromatic photometry deals with the comparison of the intensities of lights differing in colour. In the use of the photometers so far described it was assumed that lights whose intensities were being compared were the same in colour, and the light reflected from the screens was the same in tint. Even so, in actual practice it is generally found that such conditions are seldom realized and this is a very disturbing factor since the eye is a very good

judge of small differences in colour. In order to overcome such difficulties and to compare the intensities of light sources exhibiting marked differences in colour, flicker photometers have been designed. Now although no two surfaces, each illuminated by a light of different colour, can ever be said to be equally bright in the strict sense of the word, yet actual experiment shows that it is possible to obtain a reliable estimate for the ratio of the intensities of the sources if it is agreed that two surfaces are equally illuminated when, upon rapidly alternating one with the other, no sensation of flicker appears, the speed of alternation being such that the slightest change of either illumination produces a flicker.

The flicker photometers used in heterochromatic photometry owe their applicability to the fact that when two different colours are presented alternately in increasingly rapid succession to the eye, colour fusion occurs before brightness fusion. This statement means that over a small range of frequencies of alternation the field of view appears to be one definite tint in which there is a distinct flicker. This flicker disappears either if the speed of alternation is increased much beyond this stage or if the intensities of the two lights bear a certain relation to their distances from the illuminated screen. The first condition is not a criterion from which the intensities of the sources may be compared. If the lights are identical in colour, and their distances from the screen adjusted until the flicker disappears, it is found that $\frac{I_1}{r_1^2} = \frac{I_2}{r_2^2}$, where the symbols have their usual meanings. This is the equation ordinarily used in photometry. When the lights differ in colour, the speed of alternation lying within the critical range, the distances of the lights from the photometer are adjusted until the flicker disappears and the above equation is used to compare the candle-powers of the two sources.

The essential part of the Simmance-Abady Flicker Photometer is a plaster disc constructed as follows: In the upper portion of Fig. 17·10 (*a*) two cones are shown, the dotted portions having been removed. The parts ABC and DEF are then placed together to form a wedge, the shape of which can be gathered from Fig. 17·10 (*b*). The compound disc is mounted as shown in Fig. 17·10 (*c*), and may be rotated about a horizontal axis, GH, with the aid of a clockwork motor, K. M is a low-power microscope for viewing the edge of the disc as it rotates. The portion of the disc thus seen is illuminated by the sources of light whose intensities have to be compared. Let us suppose that it is illuminated by a light placed on the left-hand side of the diagram. The illuminated surface, as the disc revolves, will pass successively through the stages shown in the diagrams I to IV, Fig. 17·10 (*d*); the con-

ditions indicated are those which arise when the disc has rotated through successive right angles.

Providing that the speed of rotation is within the limits just stipulated, a flicker will be noticed. If a second light illuminates the other surface of the disc, in general, the flicker will persist, but it may be caused to disappear by adjusting the distances of the sources from the photometer. The two surfaces will be equally illuminated when this adjustment has been made,

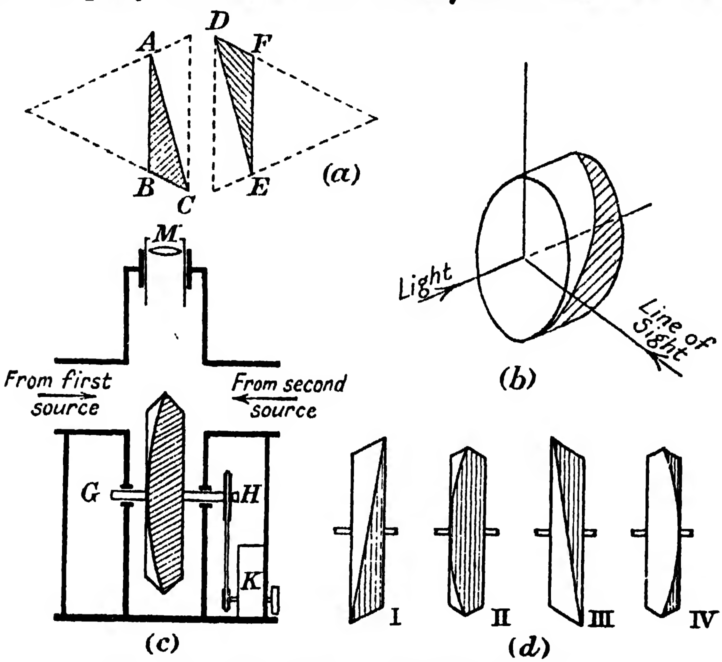


FIG. 17·10.—Simmance-Abady Flicker Photometer.

and the relative intensities of the two lights may then be calculated in the usual way.

Another flicker photometer is shown in Fig. 17·11 (*a*). AB and CD are the traces of two vertical white screens, mounted as indicated. AB is fixed, but the other screen is capable of rotation about a horizontal axis, EF, and consists of four sectors in the form of a Maltese cross. The angular width of the opaque portions equals that of those which have been removed. P_1 and P_2 are the sources of light placed so that AB is illuminated by P_1 and CD by P_2. A low-power microscope, M, is used to view a portion of the screens by the light scattered from them. If the sector is stationary the field of view is similar to that indicated in Fig. 17·11 (*b*). When the sector is made to rotate by means of the

motor provided, the field attains a uniform tint in which there is a distinct flicker at a certain speed; this is maintained and the distances of the two sources varied until the flicker vanishes. Then, with the usual notation, $\frac{I_1}{I_2} = \frac{r_1^2}{r_2^2}$.

We are justified in writing down this equation since, owing to the special shape of the disc, the candle-power of each lamp has been, in effect, reduced by the same amount.

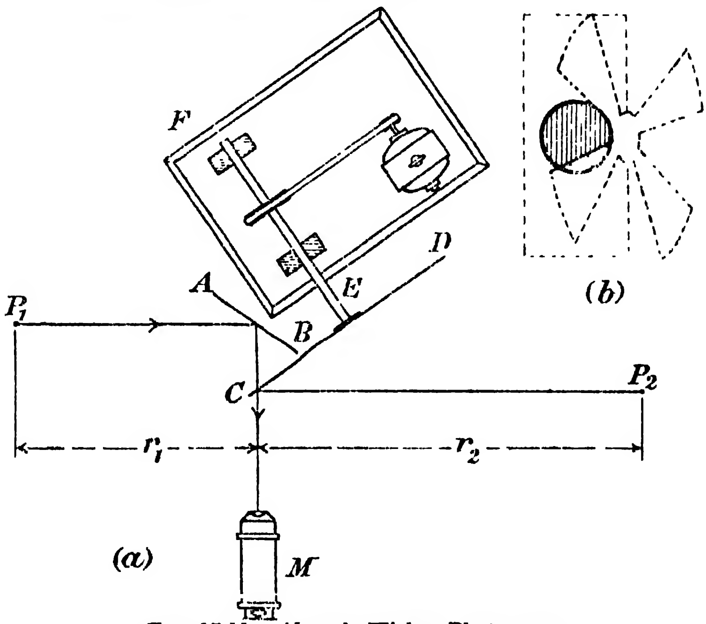


Fig. 17·11.—Abney's Flicker Photometer.

Solid Angles.—Let AB, Fig. 17·12 (*a*), be a portion of a surface and O a given point. If from O a series of straight lines are drawn to pass through points on the boundary of the area they will generate a cone—at least if they are sufficiently numerous. These lines are the so-called generators of the cone. Suppose that with O as centre a series of spheres is constructed, the above cone intercepting an area from each of them. Now the ratio obtained by dividing one of these areas by the square of the radius of the corresponding sphere is a constant for the cone OAB. From analogy with the conventional method of measuring a plane angle, the above ratio is called the ***measure of the solid angle*** subtended at O by the surface AB.

Suppose now that AB is a small area and that AC is that portion of a sphere of radius OA intercepted by the cone. If δs is the area AB, then AC = AB cos ϕ, cf. Fig. 17·12 (*b*), and, if OA = r,

$$\delta s = \text{AC} \sec \phi = r^2 \sec \phi . \delta\omega, \qquad \left[\because \frac{\text{AC}}{r^2} = \delta\omega\right.$$

where $\delta\omega$ is a measure of the solid angle OAB.

[The solid angle subtended by the surface of a sphere at its centre is 4π, since the area of the surface is $4\pi r^2$.]

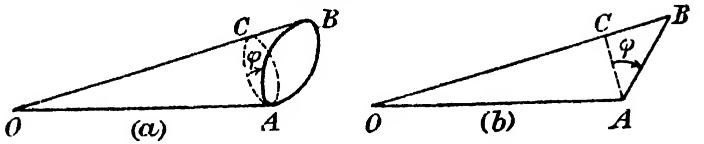


FIG. 17·12.—Solid Angles.

Further Discussion of Terms used in Photometry.—***Flux of Light.*** Just as an electric current is considered as a flow of electricity, so may light be regarded as a flow or flux of radiant energy. If the human eye were uniformly sensitive to all colours, the flux would be measured by the ***radiant power*** expressed in watts. The eye is very selective in its response to a light stimulus, so that the above method of measuring flux is not suitable and has to be replaced by an arbitrary one. In this, the flux from a luminous source is evaluated in terms of its visual effect. The unit of flux is the ***lumen,*** defined as the luminous energy emitted per second, i.e. the flux, per unit solid angle by a point source of one international candle.

It has been shown that one watt of monochromatic green light is about 620 lumens. The number of lumens associated with one watt of radiant power from a source measures the ***luminous efficiency*** of that source.

Luminous Intensity: When the source of light does not radiate uniformly in all directions, something more than a measure of its total flux is required. We refer to the ***candle-power*** or ***luminous intensity in a given direction*** of the source. Suppose that AB, Fig. 17·13 is a small area, δs, normal to the direction along which the luminou, intensity is to be measured. Let δF be the flux across this surfaces

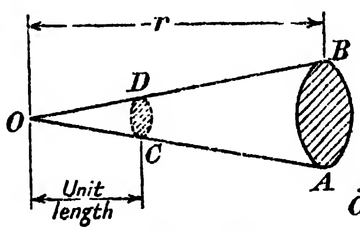


FIG. 17·13.

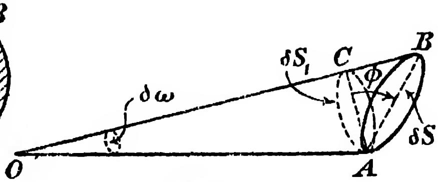


FIG. 17·14.

If r is the distance of the area from the source, and CD = $\delta\omega$, is the area cut off from a sphere of unit radius by the cone OAB, i.e. $\delta\omega$ is the measure of the solid angle OAB, then

$$\delta\omega = \frac{\delta s}{r^2}.$$

Now δF is also the flux of luminous energy in the solid angle $\delta\omega$, and the limiting value of the quantity $\frac{\delta F}{\delta\omega}$ is called the ***luminous intensity*** of the source in the direction considered, i.e.

$$I = \lim_{\delta\omega \to 0} \frac{\delta F}{\delta\omega} = \frac{dF}{d\omega}.$$

The unit of luminous intensity or candle-power is the ***international candle.*** From the above it follows that

$$\delta F = I.\delta\omega, \text{ and } F = \int dF = \int I.d\omega = I\int d\omega,$$

if the luminous intensity of the source is constant in all directions. Hence, since $\int d\omega = 4\pi$, the total flux from such a source is $4\pi I$, i.e. the flux from a point source of one candle is 4π lumens.

Illumination of a Surface. Suppose that AB, Fig. 17·14, is a small element of an illuminated surface receiving δF lumens of light, i.e. δF is the amount of light incident on AB per second. Now the luminous intensity of the source in the direction OA is $\frac{\delta F}{\delta\omega} = I$ (say). Now the amount of light falling per second on unit area of AB is

$$\frac{dF}{ds} = \lim.\frac{\delta F}{AC\sec\phi} = \lim.\frac{\delta F\cos\phi}{r^2\,\delta\omega} = \frac{I\cos\phi}{r^2} = E.$$

This expression—the fundamental relation of photometry—measures the illumination at a point on AB. The unit of illumination is the ***lux,*** defined as the illumination at a point on the surface of a sphere of radius one metre when a point source of one international candle is at the centre of the sphere.

Illumination engineers frequently evaluate the illumination of a surface in terms of the ***metre-candle.*** This is defined as the illumination produced when the light from a point source of one international candle falls on a surface one metre away from the source. Another unit is the ***foot-candle*** defined in a similar way. Although engineers use the terms metre-candle and foot-candle, each is incorrect dimensionally, for it is implied in each instance that illumination is the product of the candle-power of the source and its distance from the illuminated surface. The metre-candle is numerically equal to an illumination of one lumen per square metre and this unit is the ***lux.***

The Distribution of Light from a Given Source.—The distribution of light from an ordinary source of light varies with the direc-

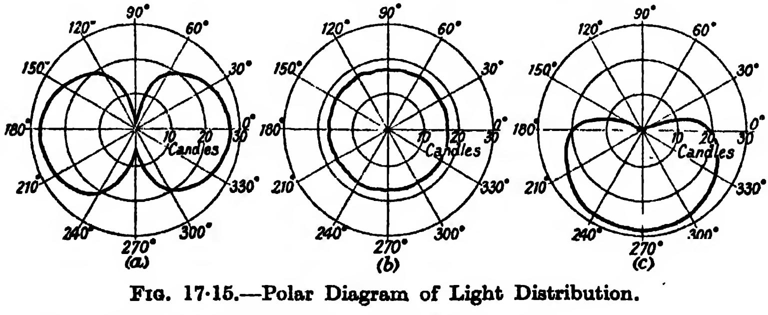


FIG. 17·15.—Polar Diagram of Light Distribution.

(*a*) In plane passing through axis of electric lamp.
(*b*) In plane normal to axis of electric lamp.
(*c*) In plane passing through axis of electric lamp, but with a reflector.

tion in which it is measured. Hence, in order to specify a light source completely, it is necessary to observe its luminous intensity in different

directions. A convenient method of indicating such distributions is by means of diagrams similar to those shown in Fig. 17·15. The curve in any instance is constructed by drawing radii vectores from a point to represent the magnitude of the luminous intensity in a direction parallel to that of the particular radius vector concerned, and then joining the extremities of these lines. The distribution of light in other planes is, in general, different from those shown.

Mean Spherical Candle-Power. Integrating Photometers.—Let I be the candle-power in a given direction of a source, Fig. 17·16 (*a*), placed at the centre of a sphere of unit radius. Then the flux of luminous energy from that source across an area $\delta s = \delta\omega$, through which the given direction passes, is given by $\delta F = I.\delta\omega$. Hence the total flux of luminous energy per second from the source is $\int I d\omega$.

Now the area of the surface of the unit sphere is 4π, so that the mean flux of luminous energy across unit area of the above sphere, i.e. the mean flux per unit solid angle, is

$$\frac{1}{4\pi}\int I\, d\omega = I_0.$$

This is termed the ***mean spherical candle-power*** of the source.

Integrating Photometers.—The total flux of luminous energy from a given source of light, and hence its mean spherical candle-power, may be computed by the laborious process of measuring the luminous intensity of the source in many different directions, and then constructing the polar curves for different planes. SUMPNER, in 1892, showed that if a source of light is placed in a large hollow

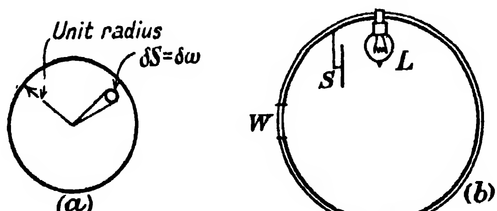


FIG. 17·16.—An Integrating Photometer.

sphere whose walls are perfectly diffusing, then the illumination at any point on the wall due to light reflected from the remainder of the walls is a constant. This principle was put into practical use by ULBRICHT in 1900. A large spherical globe is painted white (zinc oxide paint) and the lamp L, Fig. 17·16 (*b*), placed in position. The globe is provided with an opal glass window, W, screened from direct rays from L, the lamp under test, by the screen, S, also painted white on both sides. The illumination at W is then constant for a given lamp providing that the white surface is uniformly white and has no selective action on light of different wave-lengths.

The intensity of the light transmitted through W is then compared with the intensity of a constant source by some other photometer.

The lamp L is then replaced by one whose mean spherical candle-power has been determined by the method suggested above and the intensity of the light transmitted through W again compared with that of the standard source.

The ratio of the two intensities is the ratio of the mean spherical candle-powers of the two lamps which have been placed in the integrating photometer.

Photometers of this type and having a diameter of 88 inches have been made. In recent years the National Physical Laboratory at Teddington has designed a cubical photometer of this nature since it is much easier to construct. Theoretically, this whitened cube photometer is less accurate than a spherical one; in practice, however, it is very efficient.

Indoor and Street Lighting.—So far we have only discussed methods of measuring the actual illumination at a point. Now in designing any particular lighting system the degree of illumination is only one factor to be considered. Other requirements have to be met; these differ according as the lighting is for indoor or street purposes. Let us consider them in turn.

The requirements for an indoor system are (*a*) adequate illumination, (*b*) absence of glare, (*c*) non-excessive contrast, and (*d*) proper distribution. The intensity of illumination varies with different types of room and depends chiefly on the type of work to be done in them. For household purposes an illumination of 30 to 60 lux (metre-candles) is satisfactory. If no fine work is to be done, as in a foundry, the illumination may be reduced to one-third of the above values. On the other hand, the routine work of a drawing office requires an illumination of about 90 lux.

By minimizing the glare from a light source much eye-strain may be avoided.

Excessive contrasts are particularly annoying, and are a source of constant danger if they exist in street or factory lighting. On the other hand, a watchmaker finds that a complete absence of contrast is not desirable for his work.

As regards the distribution of light a concentration of light in one part of a room is to be avoided; if work is to be done at several points in a room then a general illumination of 5 lux is generally considered sufficient when at each point where the work is carried on, there is a supplementary light.

In street lighting the illumination should be as even as possible. In practice this is accomplished by using specially designed reflectors so that the light is concentrated along the surface of a cone of very wide angle, or by placing the lamp at a considerable distance above the road level.

EXAMPLES XVII

1.—In a Bunsen photometer two lamps are placed 62·7 and 84·6 cm. from the 'spot' when there is equality of contrast between the waxed and unwaxed portions on each side of the paper. Compare the candle-powers of the two lamps.

2.—Two lamps, whose candle-powers are as 2·5 : 1, are 150 cm. apart. At what distance from the less bright lamp must a grease spot be placed so that there is equality of contrast on the two sides of the 'spot'?

3.—Describe how you would determine the fractional reduction in the candle-power of an electric lamp when it is surrounded by a translucent globe. Assuming that the globe absorbs 8 per cent. of the light emitted by the lamp, calculate the ratio of the distances of the lamp from the photometer in the experiment you describe, assuming the distance of the comparison source to be the same in each instance.

4.—Describe some modern form of photometer, indicating the particular features of the instrument you describe. A candle and a glow-lamp of 36 candle-power are 1 metre apart. Where must a screen be placed on the line joining them so as to be equally illuminated on both sides?

5.—Describe some form of photometer. Discuss its accuracy and explain the principles on which its action depends. Where may a sheet of paper be placed on the line through two sources of light of candle-powers 5 and 4 respectively, and 2 metres apart, so as to be equally illuminated by each of them?

6.—A lamp of 3 candle-power is placed at a distance of 30 cm. from the grease spot of a photometer and another lamp of 6 candle-power is placed at a distance of 50 cm. on the same side of the instrument. Find where a third lamp of 10 candle-power must be placed in order that both sides of the photometer may be equally illuminated.

7.—Describe and explain the action of a 'grease spot' photometer. Explain how it could be used to determine the fraction of light transmitted by a sheet of imperfectly transparent substance. (L. '28.)

CHAPTER XVIII

THE REFLEXION OF LIGHT AT PLANE SURFACES

When light is incident upon any body the subsequent history of the light is determined by the nature of the body and its surface; one part of the light is returned into the medium in which it originally travelled—it is ***reflected***; the remaining portion enters the body and is there absorbed if the body is ***opaque***, or it is transmitted through the body if the latter is ***transparent***. If the body transmits light, but is such that an object cannot be seen distinctly through it, then the body is said to be ***translucent***.

The Reflected Rays.—The direction in which the reflected light is propagated is determined by the laws of reflexion. If the surface consists of innumerable small facets then light incident upon them is reflected from each little facet according to these laws, but there is no definite image formed, for the reflected rays go in all directions, as the facets will be orientated at random. In this connection an interesting experiment has been described by Wood. A piece of plane glass, smoked by passing it rapidly through a smoke flame, is not a good reflector of light, but if it is held between a source of light and the eye so that the light, glass, and eye are nearly in a straight line then a red image of the source is seen—the image becomes brighter and more white the more nearly the straight line condition is approached.

The Laws of Reflexion.—Let CD, Fig. 18·1, represent a plane sheet of polished metal [this is better than a silvered mirror, in which the thickness of the glass is a disturbing feature]. Let AO be the ray of light travelling towards the mirror, whilst OB is the reflected ray; then AO is the ***incident*** and OB the ***reflected*** ray.

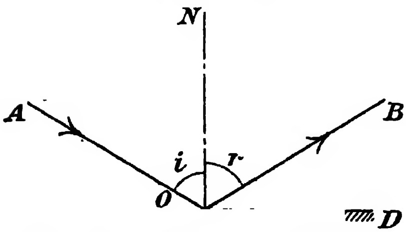


Fig. 18·1.—Reflexion of Light at Plane Surface.

Let ON be the normal (i.e. a line perpendicular to a surface) at O

to CD. Then the $\widehat{AON}$ is called the angle of incidence, whilst $\widehat{NOB}$ is the angle of reflexion. The laws of reflexion state—

(*a*) ***The incident ray, the reflected ray, and the normal to the mirror at the point of incidence lie in the same plane.***

(*b*) ***The angles of incidence and reflexion are equal.***

Experimental Verification of the Laws of Reflexion of Light Rays.—The first law of reflexion may be verified with the aid of the apparatus shown diagrammatically in Fig. 18·2 (*a*). M is a shallow pool of mercury (or black ink) which serves as a mirror which is truly horizontal, and PHK is a plumb line placed as shown. When at rest this will hang vertically downwards. O is a small object above the level of the mirror. The object may be the point at which two black lines intersect on a white card or the central point of a clear cross on an opaque base, the cross being suitably illuminated, cf. Fig. 18·2 (*b*). The eye, E, is placed in such a

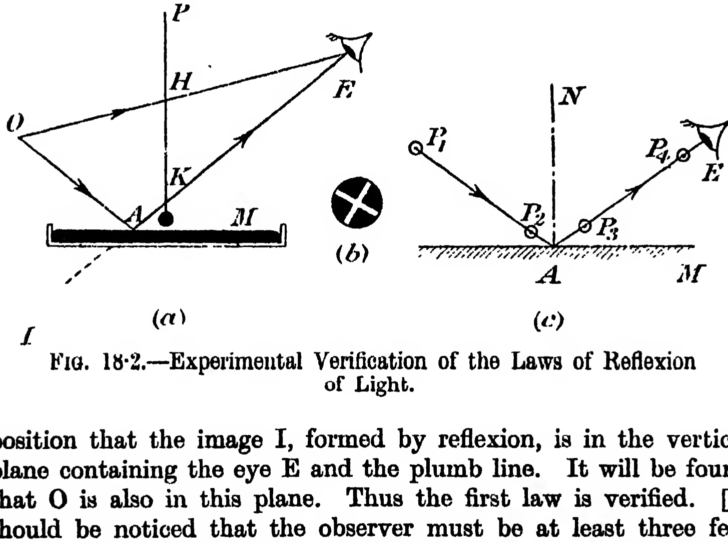


FIG. 18·2.—Experimental Verification of the Laws of Reflexion of Light.

position that the image I, formed by reflexion, is in the vertical plane containing the eye E and the plumb line. It will be found that O is also in this plane. Thus the first law is verified. [It should be noticed that the observer must be at least three feet from the plumb line and that O should be near to this line so that the line and the points O and I may be seen in focus at the same time by the observer.]

The second law of reflexion may be verified by placing a mirror M—preferably a piece of metal highly polished—so that its plane is perpendicular to a piece of paper on which it stands, cf. Fig. 18·2 (*c*). Pins P_1 and P_2 are placed to define a ray. The images of these pins formed by reflexion are observed and pins P_3 and P_4 placed in such positions that all four pins appear to be coplanar. The position of the reflecting surface is marked on the paper. The

rays P_1 P_2 and P_3 P_4 are then drawn—it is advisable for these pins to be at least 10 cm. apart so that the paths may be drawn accurately. It will be found that the rays intersect on the surface of the mirror. If the normal to the mirror at this point is drawn and the angles of incidence and of reflexion are measured, these will be found to be equal within the limits of experimental error.

Image of a Luminous Point Formed by a Plane Mirror.—An immediate consequence of the laws of reflexion is that the image of a luminous point in a plane mirror is as far behind the mirror as the source is in front and that the source and image lie on a line which is normal to the surface. This is easily proved—cf. Fig. 18·3.

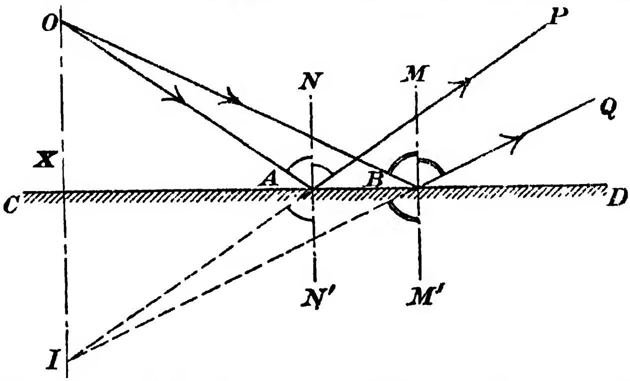


FIG. 18·3.—Image of a Luminous Point Formed by Reflexion in a Plane Mirror.

Here OA and OB represent two rays incident upon a plane mirror at A and B respectively; the reflected rays are AP and BQ. The laws of reflexion state that

(1) $O\hat{A}N = N\hat{A}P$, and $O\hat{B}M = M\hat{B}Q$ where AN and BM are normals to the mirror at A and B respectively.

(2) The rays and normals are all in one plane.

Let PA and QB when produced meet at I, so that the reflected rays apparently proceed from I. [Since they do not actually proceed from I, the lines are dotted.] If AN′ and BM′ are continuations of the normals on the other side of the mirror, then

$$\begin{aligned} O\hat{A}N &= N\hat{A}P \\ &= I\hat{A}N' \end{aligned}$$

$$\therefore\ O\hat{A}N + 90° = I\hat{A}N' + 90°,\ \text{i.e.}\ O\hat{A}B = B\hat{A}I.$$

Similarly $O\hat{B}A = A\hat{B}I$. Hence the Δ's OAB and IAB are congruent, for the base is common and the base angles of one are equal to the base angles of the second triangle [proved].

Hence in the $\triangle$'s OAX and XAI,

$$\begin{cases} AX \text{ is common} \\ OA = AI \because \triangle\text{'s} . OAB \text{ and } IAB \text{ are congruent.} \\ O\hat{A}X = X\hat{A}I \because \text{these angles are supplements of angles which} \end{cases}$$
are equal.

$$\therefore \triangle\text{'s are congruent}$$

$$\therefore OX = XI$$

$$\text{and the } O\hat{X}A = A\hat{X}I,$$

and these, being adjacent, are right angles.

Image of an Object placed in Front of a Plane Mirror.—Since an object may be regarded as a succession of points, the laws of reflexion are applicable to all the rays proceeding from every point of the body. The cone of rays which proceeds from A, Fig. 18·4, is reflected by the mirror MM, so that it apparently proceeds from A_1 the image. In drawing such a cone of rays, it is better to join the eye to the image A_1, the position of the cone behind the mirror being indicated by dotted lines. From the points where this cone meets the mirror, straight lines are drawn to A. The path of the rays of light from A to the eye is then completely known. Similarly for B and its image B_1. The image is obtained by joining A_1B_1.

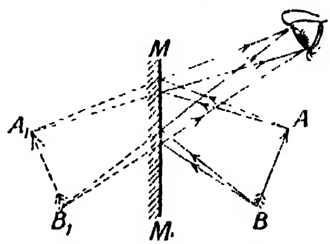


FIG. 18·4.—Image of a Finite Object Formed by Reflexion at a Plane Mirror.

Inclined Mirrors.—When two plane mirrors are inclined to each other, an object placed between them gives rise to images in both mirrors. These images *may* give rise to other images, the number of new images formed by repeated reflexion depending upon the angle between the mirrors. We shall consider two particular instances :—

(*a*) ***Two mirrors inclined at 90°.***—In Fig. 18·5, O is an object placed in front of the two mirrors, and in order to simplify the drawing only the extreme point at the top of O is considered as being luminous. The images I_1 and I_2 are formed by reflexion from M_1 and M_2 respectively. Now the image I_1 is ***in front of the plane containing*** M_2, and therefore can produce an image by reflexion in M_2; that image is I_{12}. Similarly, I_2 is in front of M_1 and gives rise to an image I_{21} coinciding with I_{12}. In order to draw the pencil or cone of light, which proceeds from O and enters the eye E, so placed that it perceives the image at I_{12}, the pencil $EI_{12}E$ is

drawn. Now I_{12} is the image of I_1, so that if this pencil cuts M_2 in *cd* the pencil of cI_1d is drawn; if this cuts M_1 in *ab* then O is joined to *a* and *b*. The parts of the three pencils thus drawn,

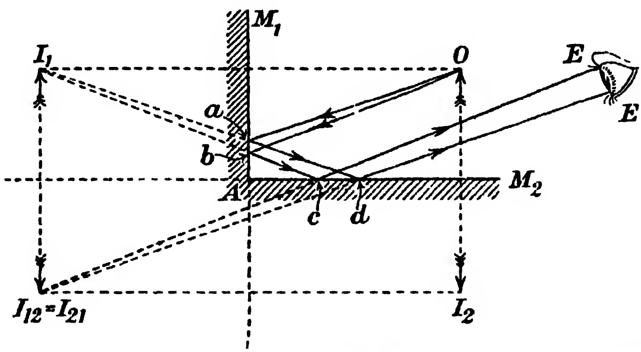


FIG. 18·5.—Formation of Multiple Images with Mirrors Inclined at 90°.

which lie between the mirrors, represent the cone of light which is required.

(*b*) ***Two mirrors inclined at 60°.***—This particular instance is illustrated in Fig. 18·6. The images, I_1 and I_2 are formed by one reflexion at each of the mirrors respectively. Light from these

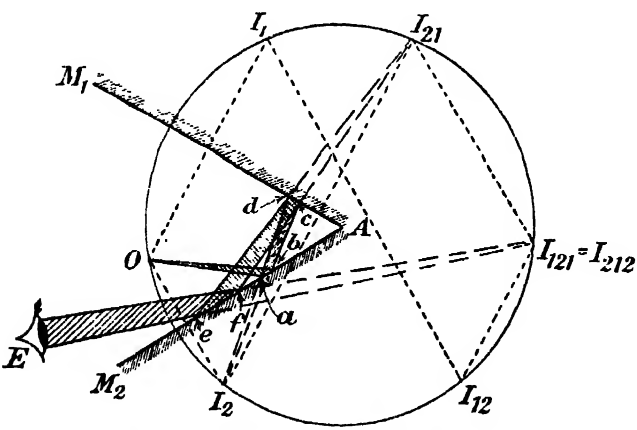


FIG. 18·6.—Formation of Multiple Images with Mirrors inclined at 60°.

images then produces two further images I_{12} and I_{21} and finally I_{121} and I_{212}, which coincide, are formed by the light which is again reflected from the mirror.

In order to draw the cone of rays which gives rise to the formation of I_{212}, let us say, the eye E is placed in some convenient position. It must be remembered that an image can only be seen when the light rays diverging from it actually enter the eye. Accordingly the cone $I_{212}E$ is drawn; it will intersect M_2 in e and f respectively, so that Eef represents the rays which enter the eye after reflexion at M_2. The image I_{212} arises as a reflexion of I_{21} in M_2, so that when the cone $edI_{21}cf$ is drawn, the portion $cdef$ represents the rays which finally give rise to I_{212} after reflexion at M_2. Now I_{21} is the image of I_2, so that, as above, the rays in the pencil $abcd$ are those which finally complete I_{212}. By joining the points a and b to O the complete cone of rays is obtained.

It will be noticed that all the images lie on a circle whose centre is A, the point of intersection of the mirrors. In both these problems, as soon as an image is formed so that it is behind *both* mirrors, no further reflexions are possible.

Parallel Mirrors.—When a luminous body is situated between two parallel mirrors, the image is always in front of one of them, so that an infinite array of images all lying on one normal to the mirrors is formed. In practice this is not so, because the light, after several reflexions, is so weakened that the retina of the eye fails to detect more than about ten images. Also, the mirrors are never exactly plane or parallel—the images, therefore, lie on the circumference of a circle whose radius is large.

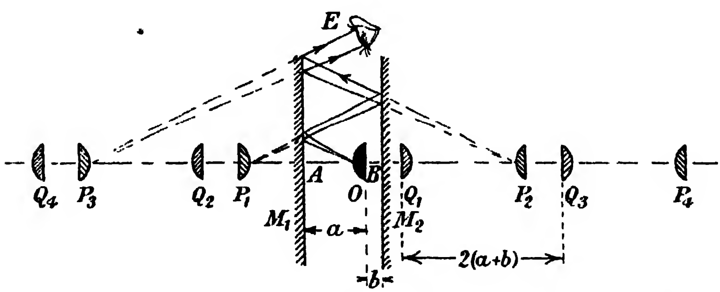


FIG. 18·7.—Multiple Images by Reflexion in Parallel Mirrors.

Let O, Fig. 18·7, be an object situated between two parallel mirrors, M_1 and M_2. In the absence of loss of light, two infinite series of images will be formed.

Thus P_1 is the image of O in M_1; P_2 that of P_1 in M_2; P_3 that of P_2 in M_1, etc. Similarly for the Q series. The paths of the rays of light from O to an eye E looking at P_3, for example, are traced in the manner indicated.

Suppose that a and b are the distances of O from M_1 and M_2 respectively. Then

$$P_1A = AO = a \qquad Q_1B = BO = b$$
$$P_2B = 2a + b \qquad Q_2A = a + 2b$$
$$P_3A = 3a + 2b \qquad Q_3B = 2a + 3b$$

Hence $P_1P_3 = P_2P_4 = P_3P_5 = \qquad = 2(a + b)$

$$Q_1Q_3 = Q_2Q_4 = Q_3Q_5 = \qquad = 2(a + b)$$

The above equations show that the distance between consecutive images of either series behind the mirrors is constant, and equal to twice the distance apart of the mirrors.

Reflexion of a Ray by a Rotating Plane Mirror.—To measure a small rotation by an optical contrivance, use is made of the fact that if a mirror is rotated through an angle θ, the reflected ray is

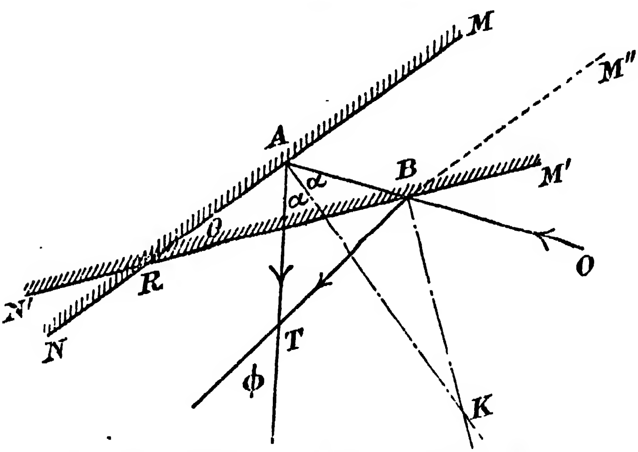


FIG. 18·8.—Reflexion by a Rotating Plane Mirror.

rotated through an angle 2θ, the incident ray following the same path both before and after the rotation of the mirror.

Proof: Let OA, Fig. 18·8, be a ray of light incident upon the mirror MN at A; let AT be the reflected ray. Suppose that the mirror is rotated through an angle θ to the position M′N′. The incident ray then meets the mirror at B, so that the reflected ray becomes BT. Let AK and BK be the normals at A and B respectively. Through B draw BM″ parallel to NM. Then

$$M'\widehat{B}O = M''\widehat{B}O - M''\widehat{B}M'$$
$$= M\widehat{A}O - M''\widehat{B}M'$$
$$= (90° - \alpha) - \theta,$$

where a and θ are the angles indicated. Hence, since the angles of incidence and reflexion are equal,

$$N'\widehat{B}T = 90° - a - \theta.$$

Also $$A\widehat{B}R = M'\widehat{B}O.$$

Let ϕ be the angle through which the ray of light rotates. Then, since the sum of the angles of a $\triangle$ is 180°,

$$\begin{aligned}\phi &= 180° - a - a - (90° - a - \theta) - (90° - a - \theta)\\ &= 2\theta.\end{aligned}$$

Hence the angle through which the ray rotates is twice the angle through which the mirror moves.

The Sextant.—This instrument was developed by HADLEY in 1731 so that sailors might ascertain their latitude by measuring the angle of elevation of a star at a stated time. He writes: 'The instrument was designed to be of use where the motion of the objects or any circumstances occasioning an unsteadiness in the common instruments render the observations difficult or uncertain.' It takes the place of the theodolite in surveying and its distinct

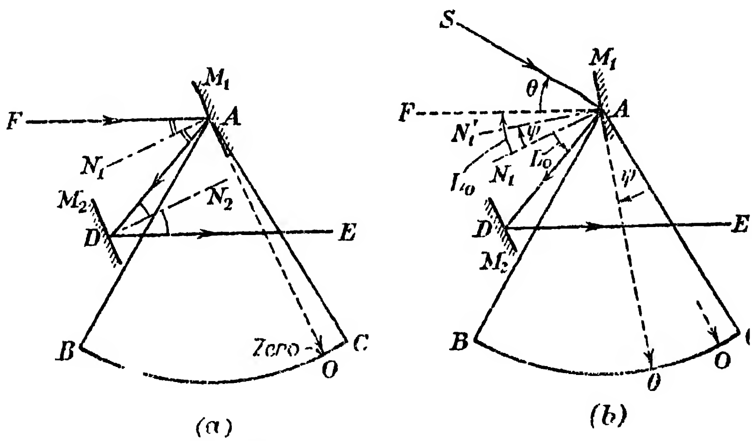


FIG. 18·9.—The Principle of a Sextant.

feature is that images of two points, whose angular separation is to be determined, are in the field of view at the same time. Before describing a sextant let us make the following digression.

Let ABC, Fig. 18·9 (a) be a framework in which $B\widehat{A}C = 60°$. M_1 is a plane mirror with its face normal to the plane of the diagram: it is capable of rotation about an axis through its mid-point A normal to ABC. M_2 is a second plane mirror with its face also normal to the plane of the diagram. Its mid-point D is not

necessarily on AB, but its trace is such that a ray travelling along AD is reflected along DE where DE is parallel to the straight line BC: the inclination of M_2 is fixed by the fact that the normal DN_2 to it at D must bisect the angle ADE. The zero position of M_1 is determined by the condition that a ray FA, parallel to DE, must be reflected by M_1 so that FA traverses in all the path FADE. If AN_1 is normal to M_1 at A, then when the above condition is satisfied,

$$F\widehat{A}N_1 = N_1\widehat{A}D\,; \quad A\widehat{D}N_2 = N_2\widehat{D}E$$

But
$$FAD = A\widehat{D}E.$$

Hence
$$N_1\widehat{A}D = A\widehat{D}N_2,$$

or the mirrors must be parallel to each other when a pointer rigidly attached to M_1 indicates 'zero.'

It is now necessary to determine the angle through which M_1 must be rotated in order that a ray SA, Fig. 5·8 (*b*), after reflexion at M_1, shall proceed along AD and hence along DE, since M_2 is fixed, the line DE being horizontal. Let $\theta = S\widehat{A}F$ and ψ the angle of rotation required. If AN_1' is the normal to M_1 at A in the new position of this mirror, $N_1\widehat{A}N_1' = \psi$. Let i_0 be $F\widehat{A}N_1 = N_1\widehat{A}D$. Then $S\widehat{A}N_1' = N_1'\widehat{A}D$,

or
$$\theta + i_0 - \psi = i_0 + \psi.$$
$$\therefore \psi = \tfrac{1}{2}\theta.$$

Accordingly, the pointer attached to M_1 moves over a circular scale on BC, the graduations having nominal values equal to twice their real value so that the angle of elevation required may be read off directly.

An arrangement of plane mirrors similar to the above is used in the modern sextant, shown in Fig. 18·10 (*a*). ABC is the metal framework, the angle BAC being 60°. [In Hadley's original instrument this angle was 45°.] M_1 and M_2 are the plane mirrors mounted with their reflecting faces normal to the plane ABC, M_2 being fixed to the framework, while M_1 is fixed to an arm, L, capable of rotating about an axis through A normal to the plane of the diagram. The arm L carries a vernier, the arc BC is graduated as before and, as shown previously, M_1 is parallel to M_2 when the zero on the vernier scale coincides with the zero on BC. When the plane of M_1 lies along AC, a ray, FA, parallel to DE, will traverse the path FADE. If θ is the angle SAF, the angle between the positions occupied by M_1 in the two instances is $\tfrac{1}{2}\theta$. The position of M_1 is indicated by a pointer attached to M_1, and moving over an angular scale on BC.

T is an erecting telescope [Galilean type—cf. p. 519] whose axis is parallel to BC and intersects M_2 in D. The inclination of M_2 is determined as above, i.e. its face is perpendicular to DN_2, the bisector of the angle ADE.

Let us assume that the above instrument is to be used to determine the angular elevation of the top of a spire with respect to the top of a wall directly below it. The telescope T is directed towards the top of the wall and a view is possible since only one half of M_2 is silvered. The arm L is rotated until the image of the wall-top seen by reflexions at M_1 and M_2 appears to coincide

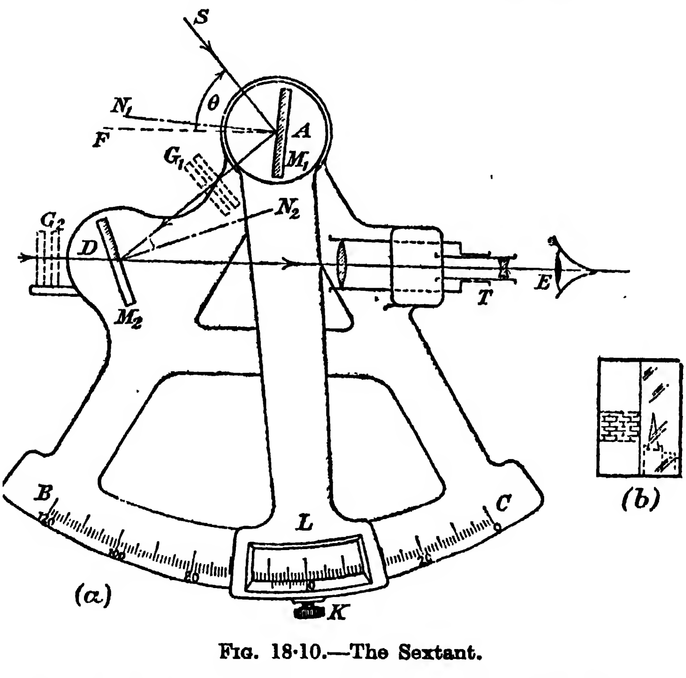


Fig. 18·10.—The Sextant.

with the image seen directly. The arm L is clamped by means of the screw K. The scale reading of the vernier is noted. The arm L is then rotated until an image of the top of the spire, formed after reflexion at M_1 and M_2, appears to coincide with the image of the wall-top seen directly—cf. Fig. 18·10 (*b*). The vernier reading is again noted. The difference between these readings gives the angular elevation required.

G_1 and G_2 are pieces of tinted glass which are placed in the paths of the different light rays if these proceed from a brilliant source: they *must* be used whenever an attempt is made to measure the altitude of the sun.

It should be noted that the eyepiece of the telescope is not provided with cross-wires, since the telescope is only used for examining a coincidence between two images.

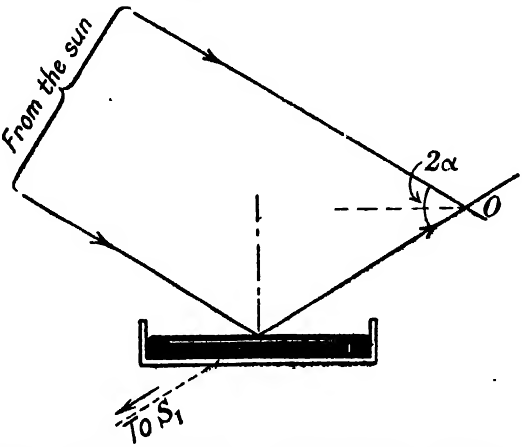


FIG. 18·11.—Artificial Horizon: Determination of Sun's Altitude.

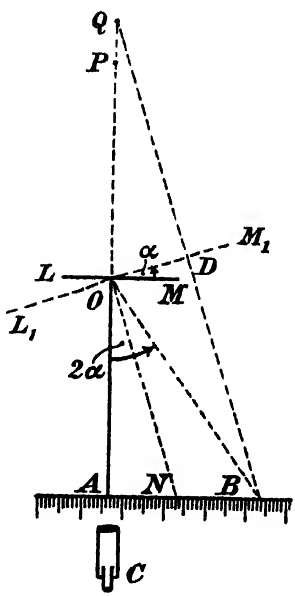


FIG. 18·12.—Measurement of Small Angular Deflexions.

To determine the Sun's Altitude.—A large dish containing mercury is placed in a convenient position and the image, S_1, Fig. 18·11, of the sun in the mercury observed (dark glasses must be used). [If a plane mirror is used it must be carefully levelled.] Suppose the sextant is at O. To determine the zero error of the instrument, an image of S_1 reflected from M_1 and M_2 [Fig. 18·10 (*a*)], coincides with the image of S_1 as seen through the clear portion of M_2. Let the vernier reading be θ_1. The mirror M_1 is then rotated until an image of the sun reflected in the mirrors M_1 and M_2 coincides with S_1. Let the vernier reading be θ_2. If α is the sun's altitude, $2\alpha = \theta_2 - \theta_1$.

The Measurement of Small Deflexions.—When a body rotates through a small angle about a vertical axis, that angle may be measured by a method due to POGGENDORFF. A telescope, C, Fig. 18·12, is

placed normally to a small plane mirror rigidly attached to the moving system. A scale, graduated in cm., etc., is placed near to the telescope and parallel to the rest position of the mirror, LM. An image of a point A on the scale will be seen through the telescope, the image being at P, where OA = OP. When the mirror moves to a position L_1M_1 the image of another point, B, on the scale will be seen at Q, where BD = DQ and BD is normal to L_1M_1. Let a be the angle through which the mirror has turned; then the $\widehat{AOB}$ is $2a$, and $\tan 2a = \frac{AB}{OA}$

For small angles $\tan 2a = 2a$, so that $a = \frac{AB}{2 . OA}$.

EXAMPLES XVIII

1.—A ray of light is reflected from a plane mirror after incidence at an angle of 43·5°. The mirror is rotated through 31°. Find, by drawing, the angle through which the reflected ray is rotated.

2.—Show that when a ray of light is reflected from a plane mirror it travels along the shortest possible path.

3.—Two mirrors intersect at right angles. Prove that, if a ray of light is reflected from both mirrors, the emergent ray will be parallel to its original direction.

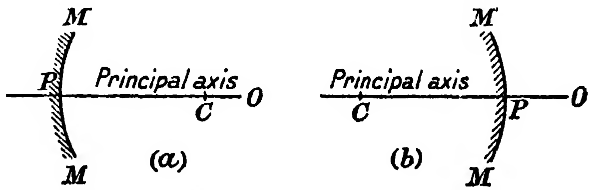
4.—State the laws of reflexion of light. A small object is placed between two plane mirrors inclined at an angle of 60°. Determine graphically the number of images formed, and indicate the path of the rays when an image formed by two reflexions is observed.

5.—State the laws of reflexion of light. Two plane mirrors are inclined at a fixed angle to one another and the combination can be rotated about their line of intersection as axis. Show that, if a ray of light is reflected first in one mirror and then in the other in a plane at right angles to the axis, the deviation of the ray is unaltered by the rotation of the mirrors.

CHAPTER XIX

REFLEXION OF LIGHT AT SPHERICAL SURFACES

Preliminary Definitions.—A polished surface having the form of a portion of a sphere is termed a spherical mirror. The centre of curvature of the mirror is the centre of the sphere, C, Fig. 19·1. If the inside of the spherical cap acts as a mirror, the reflecting surface is said to be *concave*; if the outside reflects, it is a *convex* mirror. Suppose that O is a point source of light and that a



A Concave Reflecting Surface. A Convex Reflecting Surface.

FIG. 19·1.

straight line through O and C (produced if necessary) cuts the surface of the mirror in P. Then P is termed the *pole of the mirror*, while the line CP is termed the *principal axis*. The boundary or periphery of a mirror is usually circular, and the length of the diameter of this circle is called the *aperture*. The radius of curvature of a spherical mirror is the radius of the sphere of which the mirror forms a part.

Some Optical Conventions.—Distances are measured along the axis from P, the pole of the mirror; in this book distances measured to the right of P are considered positive; those to the left are negative. Since some prefer to call positive those distances which are measured against the direction in which light travels, and negative distances measured in the opposite direction, we shall always place our object on the right of the mirror or lens so that readers of this book may use either convention.

Reflexion at a Concave Spherical Surface.—The principal section of a concave spherical surface is shown in Fig. 19·2. Let O be a luminous point on the axis of this concave reflecting surface

—or mirror; let OA be a ray incident at A. Then the line CA, which is a radius, is normal to the surface at A. The reflected ray is therefore AI, where $\widehat{IAC} = \widehat{OAC}$. Now OC is a ray of light which travels along a radius; it is therefore reflected along this radius in the reverse direction—the two reflected light rays meet at I, and this will be the image of O if it can be shown that all the reflected rays pass through this point. If so, the image will be ***real*** because the rays of light ***actually*** pass through it, i.e. it may be obtained on a screen.

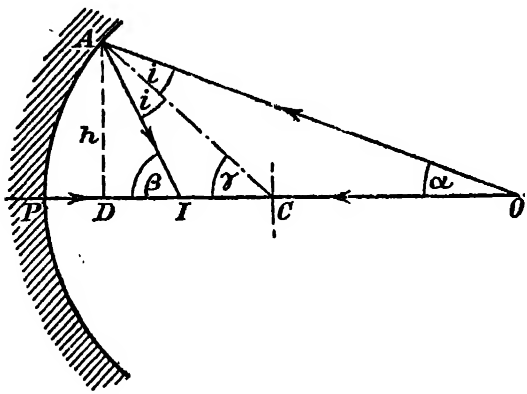


FIG. 19·2.—Reflexion at a Concave Spherical Surface.

Let the points O and I be at distances u and v respectively from P, while r is the radius PC. Let i be the angle of incidence at A—it is therefore the measure of the angle of reflexion also; let α, β and γ denote the angles shown. Since the exterior angle of a triangle is equal to the sum of the two interior and opposite angles, we have

$$\gamma = i + \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\beta = 2i + \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Eliminating i from these equations, by multiplying (1) by 2 and then subtracting, we obtain

$$2\gamma - \beta = \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The assumption is now made that the angles α, β and γ are small, so that their circular measure is expressed by their tangents, i.e. the theory developed here is only applicable to ***paraxial rays,*** i.e. rays near to the principal axis of an optical system. Draw AD perpendicular to OP and call this length h.

Then $\alpha = \dfrac{h}{DO} = \dfrac{h}{PO} = \dfrac{h}{u}$, $\because$ DO $\simeq$ PO[1], since A and P are close together.

[1] The sign $\simeq$ means 'is approximately equal to.'

Similarly, $$\beta = \frac{h}{v}, \text{ and } \gamma = \frac{h}{r}.$$

Substituting these values in (3) and dividing by h throughout, we obtain

$$\frac{2}{r} - \frac{1}{v} = \frac{1}{u},$$

or

$$\frac{1}{v} + \frac{1}{u} = \frac{2}{r} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (4)$$

Let us now suppose that the object is at an infinite distance from the mirror. The particular point at which the image is then formed is called the ***focus*** of the mirror—the distance of this point from P is called the ***focal length*** (f) of the mirror, i.e. $f = [v]_{u\to\infty}$, where this symbol denotes the particular value of the image distance when $u \to \infty$. Hence substituting in (4) we obtain

$$\frac{1}{[v]_{u\to\infty}} + \frac{1}{\infty} = \frac{2}{r}, \text{ or } \frac{1}{[v]_{u\to\infty}} = \frac{2}{r},$$

i.e.

$$[v]_{u\to\infty} = f = \frac{2}{r},$$

i.e. the focal length of the mirror is one-half its radius of curvature. The formula (4) may therefore be written

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

Reflexion at a Convex Spherical Surface.—The principal section of a convex spherical surface is shown in Fig. 19·3. OA is

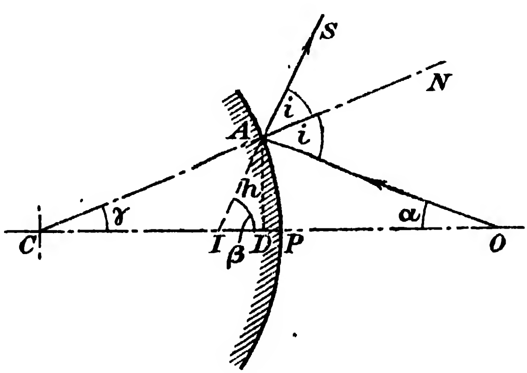


Fig. 19·3.—Reflexion at a Convex Spherical Surface.

a ray from the object O, incident at A, and reflected along AS. CAN is the normal at A. Let α, β, γ, and i be the angles as shown.

Then, as before,

$$2i = \alpha + \beta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

$$i = \gamma + \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

so that

$$0 = (\alpha + \beta) - (2\gamma + 2\alpha)$$
$$= \beta - 2\gamma - \alpha$$

$$\text{or } \alpha - \beta = -2\gamma \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

If these angles are small, then, as before,

$$\alpha = \frac{h}{u}, \ \beta = -\frac{h}{v}, \ \gamma = -\frac{h}{r},$$

the negative sign being prefixed in order to obtain positive values for the angles β and γ. Under these conditions

$$\frac{1}{v} + \frac{1}{u} = \frac{2}{r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

The image in this instance is ***virtual,*** since the rays of light do not actually pass through I, i.e. no image is obtained on a screen placed at I. Again if f is the value of v when $u = \infty, f = [v]_{u \to \infty} = \frac{r}{2}$. Hence the equation

$$\frac{1}{v} + \frac{1}{u} = \frac{1}{f} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

represents the relationship between u, v, and f for both classes of spherical reflecting surfaces, when due regard is paid to signs.

Object of Finite Dimensions.—Hitherto the object has been supposed to be a luminous point; it is now necessary to determine the nature of the image when the object is finite [but small, in order to comply with the conditions which have been stipulated]. This is best done graphically. Thus, let OA, Fig. 19·4 (*a*), be a small object placed in front of and perpendicular to the principal axis of a concave mirror whose pole is at P. Through P draw PD normal to PO to represent [on a large scale] the small element of the reflecting surface near P. PD is the trace, in the plane of the diagram, of the ***principal plane*** of the mirror. To remind us that the mirror is a concave one two arcs of circles are constructed as shown, the polished surface being indicated by the heavier line.[1] The ray AD parallel to OP passes through the principal focus F after reflexion. The ray AC passing through the centre of curvature of the mirror travels along a normal to it and therefore retraces its path after reflexion. If these two rays are drawn, they intersect

[1] *N.B.*—Students, whenever the occasion to work examples in optics or to construct diagrams of optical systems arises, are advised always to state the sign convention they use, and also how they indicate a curved mirror or lens on a diagram.

in B. This represents the image of A. If BI is drawn perpendicular to PO, then IB indicates the image.

To trace the paths of the rays by means of which an eye placed near to the principal axis sees the image, let EG be the pupil of the eye. Then all the rays proceeding from I to the eye must lie in the pencil BGE. If GB and EB are produced to meet the mirror in *g* and

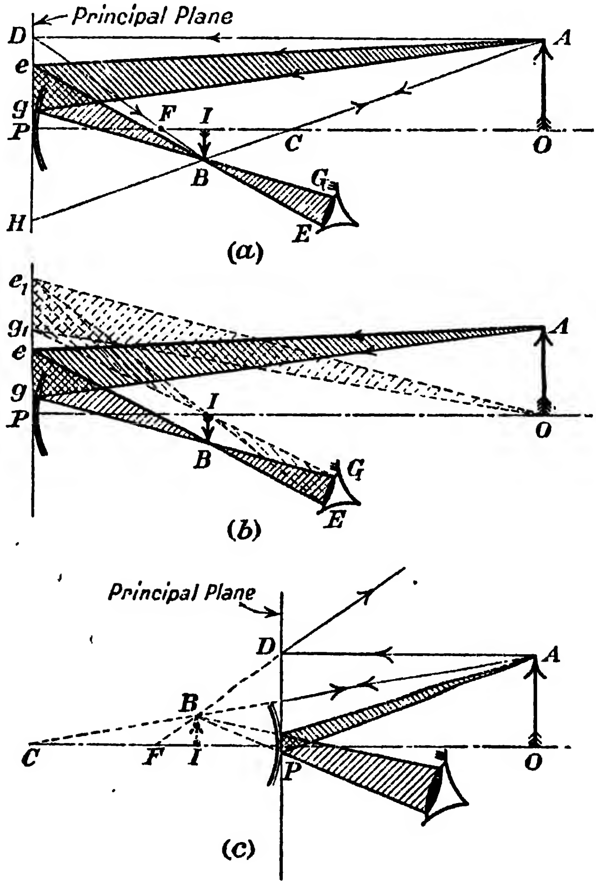


Fig. 19·4.—Formation of Images in Spherical Mirrors.

e respectively and these points joined to A we have paths of the extreme rays of the pencil by which an observer views the point B in the image.

To prevent the diagram from becoming unduly complicated the positions of the image and object have been redrawn in Fig. 19·4 (*b*). The rays A*e*E and A*g*G are then constructed as described above. To

determine the confines of the pencil of rays by which the point I in the image is seen, the lines GI and EI are produced to cut the mirror in g_1 and e_1, which points are then joined to O.

The method of determining the position and characteristics of the image formed by the reflexion of light at a convex surface is illustrated in Fig. 19·4 (*c*). Since, in this instance, the image is produced behind the surface, dotted lines are used to indicate the apparent paths of the rays in this region. The paths of the extreme rays of the pencil by which an observer sees the point B in the image are obtained as before, and it is left as an exercise for the student to draw the rays from O to the eye.

The Tracing of any Ray of Light after being reflected at a Spherical Surface.—Let OA, Fig. 19·5 (*a*) be a ray of light incident upon a concave mirror with pole, focus, and centre of curvature as shown. A plane through F and normal to the principal axis of the mirror is termed the ***focal plane*** of the mirror. Through C draw CH parallel to OA to cut the focal plane in K. Then OAHC may be considered as a beam of parallel light falling on the mirror: it will be brought to a focus in the focal plane. Since CH

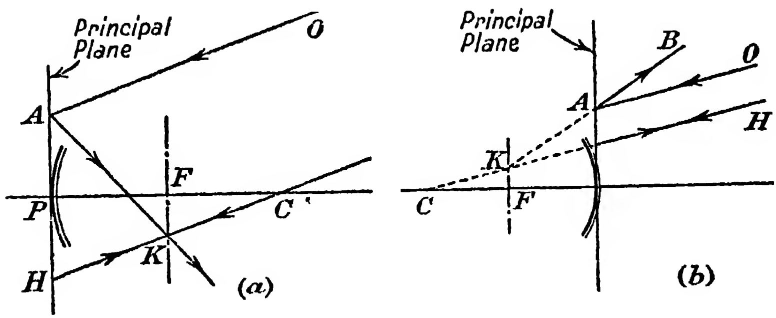


FIG. 19·5.

is one ray of the beam and it is reflected without deviation so that the reflected ray cuts the focal plane in K, it follows that K is the focus for the parallel beam considered. Hence AK must be the path of the ray OA after reflexion at the surface.

When the mirror is convex the construction is carried out as in Fig. 19·5 (*b*). No details of the method are given since the lettering on the diagram is such that the explanation given above applies at once—it must be remembered, however, that KA is a 'virtual' ray.

Magnification.—The linear magnification, *m*, produced by an optical instrument, when the object and consequently the image, are perpendicular to the principal axis of the instrument, is defined

as the ratio of the linear dimensions of the image to those of the object. Very frequently distances above the principal axis are taken as positive, and those below as negative, so that the sign of the magnification depends upon whether the image is erect or inverted. Rules are then given by means of which one may ascertain the nature of the image, e.g. whether it is real or virtual, erect or inverted. Since all this information may be obtained from an accurately drawn diagram and such diagrams should always be made whenever a problem is attempted, we shall regard all distances above or below the axis as positive, i.e. the magnification will always be considered positive; to be consistent with this, distances to the right or left of P must be considered positive, i.e. all distances are measured by their numerical magnitudes alone when dealing with magnification.

Thus, in Fig. 19·4, the magnification is given by

$$m = \frac{\text{size of image}}{\text{size of object}} = \frac{\text{IB}}{\text{OA}}.$$

But tan OPA is $\frac{\text{OA}}{\text{PO}}$, and tan IPB is $\frac{\text{IB}}{\text{PI}}$, and since these angles are equal,

$$\frac{\text{OA}}{\text{PO}} = \frac{\text{IB}}{\text{PI}}, \text{ or } \frac{\text{IB}}{\text{OA}} = \frac{\text{PI}}{\text{PO}},$$

i.e.

$$m = \frac{|v|}{|u|},$$

where the symbols $|v|$ and $|u|$ denote the numerical values of v and u respectively.

Worked Examples:

(i) A concave mirror has a focal length of 15 cm. Find the position, size, and nature of an object 4 cm. high placed (*a*) 20 cm. (*b*) 10 cm. from the mirror.

(*a*) Since the mirror is concave, f is positive: u is also positive. Hence, inserting the numerical values for f and u in the formula $\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$, and inserting the appropriate signs only when such substitutions are made, we have,

$$\frac{1}{v} + \frac{1}{20} = \frac{1}{15}. \qquad \therefore v = 60 \text{ cm.}$$

Also $m = \left|\frac{v}{u}\right| = 3$, and an accurately drawn diagram shows that the image is real and inverted.

(*b*)
$$\frac{1}{v} + \frac{1}{10} = \frac{1}{15}. \qquad \therefore v = -30 \text{ cm.}$$

Also $m = \left|\frac{30}{20}\right| = 1{\cdot}5$; the image is virtual and erect.

(ii) A concave mirror has a radius of curvature of 10 cm. Where must it be placed so that an eye shall see an image of itself magnified 4 times ?

Since $|m| = 4$, we have $|v| = |4u|$.

Let us try $v = 4u$, i.e. v is positive, since u is positive. Then

$$\frac{1}{4u} + \frac{1}{u} = \frac{2}{10}. \qquad \therefore\ u = 6{\cdot}25 \text{ cm. and } v = 25{\cdot}0 \text{ cm.}$$

This is an impossible solution, since the final image would be formed behind the eye itself.

Now try $-v = 4u$, i.e. v is negative since u is positive. Then

$$-\frac{1}{4u} + \frac{1}{u} = \frac{2}{10}. \qquad \therefore\ u = 3{\cdot}75 \text{ cm., } v = -15{\cdot}0 \text{ cm.}$$

This is the solution required, for the image is visible to the eye.

(iii) A convex mirror has a radius of curvature of 12 cm. Calculate the position of a point object 18 cm. in front of the mirror.

We have
$$\frac{1}{v} + \frac{1}{18} = -\frac{2}{12} = -\frac{1}{6},$$

the minus sign being inserted since r is negative.

$$\therefore\ v = -4{\cdot}5 \text{ cm.}$$

i.e. the image is behind the mirror.

(iv) A man holds half-way between his eye and a convex spherical mirror, 5 ft. from his eye, two fine parallel wires so that they may be seen directly and by reflexion in the mirror. If the apparent distance apart of the wires as seen directly is five times what it is seen by reflexion, calculate a value for the radius of curvature of the mirror.

The apparent size of an object or image is measured by the angle it subtends at the eye of the observer. Let x and y be the actual sizes of the object and image respectively: the foot is taken as the unit of length. Then, with the usual notation,

$$\frac{|x|}{2{\cdot}5} \div \frac{|y|}{5 + |v|} = 5, \qquad \text{(i)}$$

$$-\frac{1}{|v|} + \frac{2}{5} = -\frac{2}{|r|}, \qquad \text{(ii)}$$

and
$$\frac{|x|}{|y|} = \frac{2{\cdot}5}{|v|}. \qquad \text{(iii)}$$

Eliminating $|x|$ and $|y|$ from (i) and (iii) we have

$$\frac{2{\cdot}5}{|v|} \times \frac{5 + |v|}{2{\cdot}5} = 5$$

$$\therefore\ |v| = \frac{5}{4} \text{ ft.}$$

Hence, from (ii)

$$|r| = 5 \text{ ft.}$$

The ‘Phantom Bouquet.’—If a small flower, OA, is placed in an inverted position in front of a concave mirror, $u > f$ and $< r$, as in Fig. 19·6, a screen preventing an observer from seeing the

flower directly, its image will be formed at IB. This will be seen when the observer's eye is near to the axis of the mirror, but the

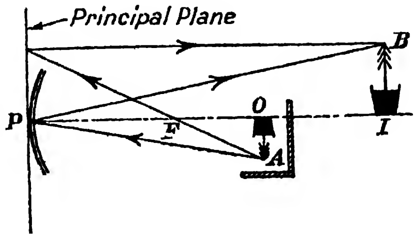


FIG. 19·6.—Phantom Bouquet.

illusion disappears when the observer steps to one side. The position of the image has been obtained by constructing the paths of the rays in the usual way.

A Particular Instance of Reflexion at a Concave Surface.—This occurs when the object is 'at the centre of curvature of the mirror'; this expression really implies that the plane containing the object is normal to the axis of the mirror, and some point in the object passes through its centre of curvature. To determine the position of the image we note that the ray OA, parallel to the axis CP, Fig. 19·7, passes through F after reflexion, and that the ray OF is parallel to PC after reflexion. The diagram shows that the magnification is unity and that the image is inverted.

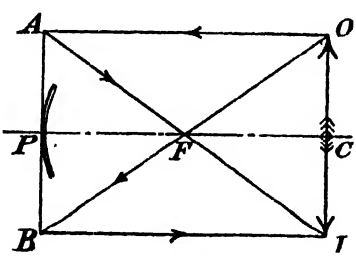


FIG. 19·7.

On the Image of the Moon in the Focal Plane of a Concave Mirror.—Let F, Fig. 19·8 (*a*), be the principal focus of a concave mirror M whose principal axis PF is directed towards the centre of the moon. Then parallel rays such as O_1Q_1 and O_2Q_2 from the centre of the moon will be focused at F after reflexion at M. Now let A_1B_1 and A_2B_2 be two parallel rays from a point on the outer edge of the moon and in the plane of the diagram. To determine the paths of these rays after reflexion at the mirror draw FK parallel to A_1B_1. Then the ray FK will be reflected along KG, parallel to PF, to cut the focal plane in G. Then G is the point in the focal plane of the mirror to which all rays parallel to FK will be focused: hence B_1G and B_2G are the reflected rays required:

FG will be the radius of the image of the moon. The diagram shows that the size of the image is determined by the focal length of the

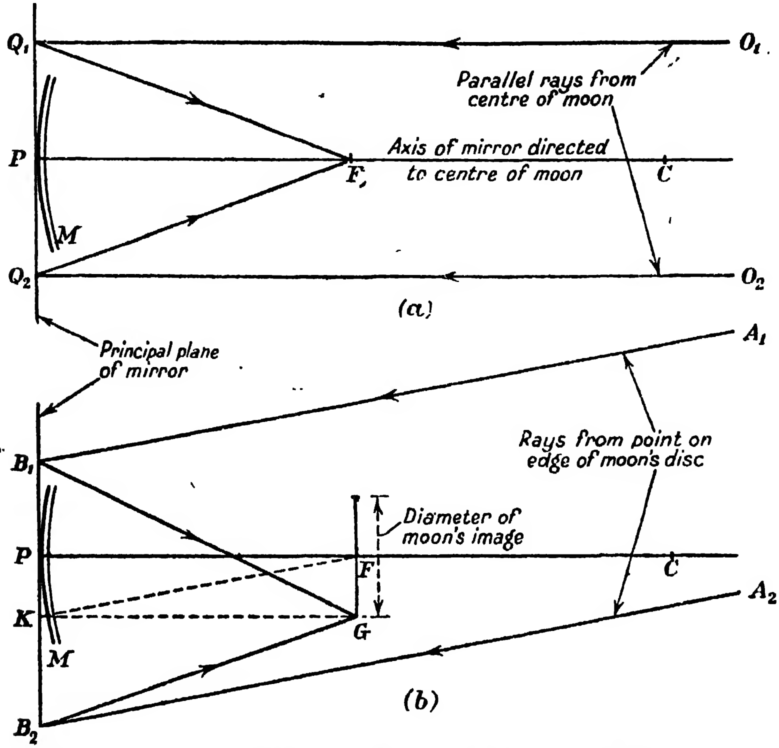


FIG. 19·8.—The Formation by Reflexion in a Concave Mirror of the Image of the Moon.

mirror: it is independent of the aperture, but the greater the aperture the brighter the image will be since the width of the beams caught by the mirror will be increased.

Spherical Aberration.—The laws of reflexion have been applied to spherical surfaces on the assumption that all the rays of light concerned were near to the optical axis, i.e. only the region of the surface in the immediate vicinity of the pole has been considered. Such a limitation was not necessary when considering the reflexion of light from plane surfaces because in such instances the image is always a perfect reproduction of the object, whereas the images produced by reflexion in curved surfaces are distorted, the amount of distortion depending upon the aperture of the mirror. Such mirrors are said to possess ***spherical aberration.***

Caustic Curve by Reflexion at a Concave Surface.—Let APB, Fig. 19·9, be the principal section of a hemispherical concave

mirror; C is the centre of curvature, F the focus, and P the pole; PC is therefore the principal axis. Suppose O to be a luminous point on the axis. The path of a reflected ray is very easily constructed because the incident and reflected rays are equally inclined to the normal [i.e. the radius] at the point of incidence. The diagram shows that all the rays reflected from points near the axis tend to pass through one point I on the axis—this is the image as

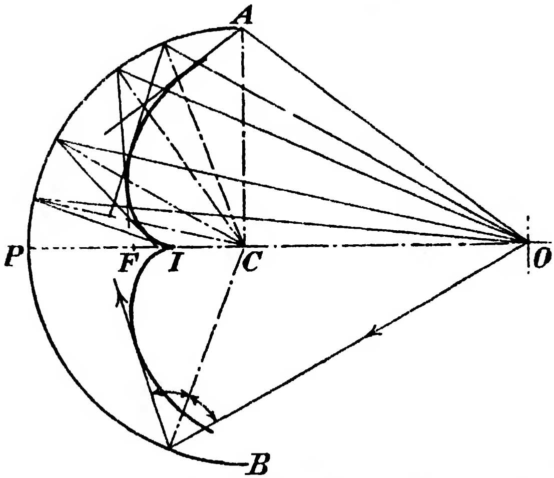


Fig. 19·9.—Caustic Curve by Reflexion at a Curved Surface.

hitherto contemplated. As the incident rays approach the direction OA, however, the reflected rays tend to cut the axis at points nearer to the mirror. If a sufficient number of reflected rays is constructed it will be found that a smooth curve can be drawn such that every reflected ray is a tangent to the curve. This is termed the *caustic curve* by reflexion at a curved surface. Such a curve is very frequently seen on the surface of tea in a cup when there is a light not directly overhead.

Focal Lines by Reflexion at a Spherical Surface.—In Fig. 19·10 a narrow pencil of rays is shown incident upon a small portion AB of a spherical surface. The two extreme rays OA and OB intersect after reflexion at F_1, a point in the plane of the paper. These same two reflected rays will cut the axis at points separated by a short distance F_2. If we imagine the figure to rotate through a small angle about the axis OC, the point F_1 will move through a short distance perpendicular to the plane of the paper, while F_2 still remains on the axis. Fig. 19·11 will perhaps help to make this clear. The lines F_1 and F_2 are termed the *first* and *second focal lines* respectively. Somewhere

between these two focal lines the reflected cone passes through a circle at right angles to the direction of propagation. This

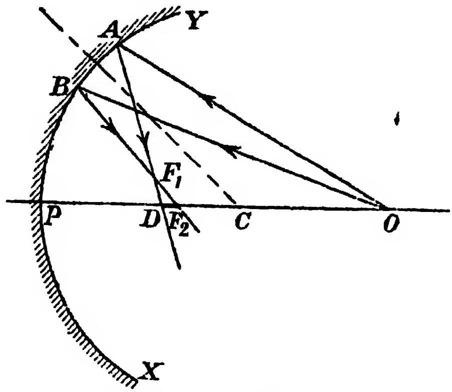


FIG. 19·10.—Focal Lines by Reflexion at a Concave Spherical Surface.

circle, known as the ***circle of least confusion,*** must exist because the width of the pencil gradually changes—at F_1 it is elongated along the perpendicular to the plane of the paper, while at F_2 it is elongated in the plane of the paper. Somewhere in between it must be equally wide in two directions—this is the region of the circle of least confusion. Such a pencil as this, which nowhere passes through a point, is termed an ***astigmatic pencil.***

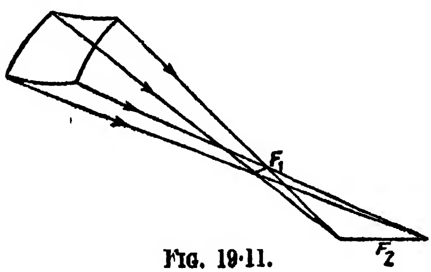


FIG. 19·11.

Parabolic Mirrors.—The parabola is a curve possessing the property that the normal at any point on it makes equal angles with a line through that point parallel to the axis, and with the line joining it to the focus. In consequence of this if a small source of light is placed at the focus, F, Fig. 19·12, of a parabola all the reflected rays will be parallel to the axis. Such mirrors are used in search-lights.

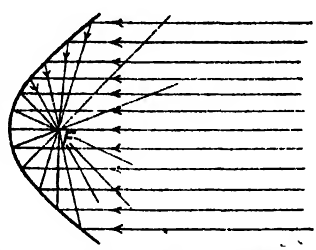


FIG. 19·12.—Parabolic Mirror.

EXAMPLES XIX

1.—An object is placed 24 cm. in front of a concave mirror when an image is formed 8 cm. from the mirror. Calculate the radius of curvature of the mirror. Check by a drawing.

2.—A candle is placed 50·3 cm. in front of a convex mirror whose focal length is 14·6 cm. Where is the image? What is the magnification? If the distance of the candle from the mirror is halved, show that the magnification is not altered in the same ratio.

3.—A concave mirror has a radius of curvature equal to 4 ft. Where must an object be placed so that the image may be magnified 3 times?

4.—Two mirrors, one convex and the other concave, each have a focal length equal to 5 in. Their poles are 18 in. apart. If an object is placed 1 ft. from the concave mirror, find the position of the image formed first by reflexion at the concave mirror, and then at the convex mirror. Check by a diagram.

5.—Establish the formula $\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$, for a convex spherical mirror. A small object is situated 8 in. in front of such a mirror having a radius of curvature equal to 6 in. Calculate the position of the image and show, on an accurately drawn diagram, the paths of the rays by means of which an eye, placed near to the axis of the mirror, sees the image.

6.—An object, 3 cm. high, is placed perpendicularly to the principal axis of a convex mirror whose focal length is 8 cm. If the object is 15 cm. away from the mirror, calculate the position of the image. Indicate, on a diagram, the paths of rays which enable an observer to see the image.

7.—A luminous object is placed 30 cm. from the surface of a convex mirror, and a plane mirror is set so that the images formed in the two mirrors lie adjacent to each other in the same plane. If the plane mirror is then 22 cm. from the object, what is the radius of curvature of the convex mirror?—(N.H.S.C. '29.)

8.—A man holds half-way between his eye and a convex spherical mirror, 4 feet from his eye, two fine parallel wires so that they may be seen directly and by reflexion in the mirror. If the apparent distance apart of the wires as seen directly is six times what it is seen by reflexion, calculate a value for the radius of curvature of the mirror.

Show, on a diagram drawn to scale, the paths of the rays of light by which an eye, near to the principal axis of the mirror, sees the image of one of the wires.

CHAPTER XX

REFRACTION AT PLANE SURFACES

The Refraction of Light.—When a ray of light passes from one homogeneous medium into another it is generally propagated in a direction which is not the same as that in which it originally travelled. EUCLID had noticed that when a ring was placed at the bottom of a vessel, it was possible to see the ring when the vessel was filled with water, even when it was impossible to see the ring in the absence of the water. He explained this phenomenon by supposing that the light from the object was refracted at the surface of the liquid.

The Laws of Refraction for Isotropic Media.—Although many illustrious workers endeavoured to discover these laws, it was not until 1621 that they were formulated by WILLIAM SNELL.

(*a*) When a ray of light passes from one homogeneous medium into another the incident ray, the refracted ray, and the normal to the surface of separation of the two media at the point of incidence are in one plane, the incident and the refracted rays being on opposite sides of the normal.

(*b*) If i is the angle of incidence, and r the angle of refraction (i.e. the angle between the normal at the point of incidence and the refracted ray) then

$$\frac{\sin i}{\sin r} = \text{constant (for light of a definite colour).}$$

This constant is called the absolute *index of refraction* of the second medium if the medium in which the light is initially travelling is a vacuum. It is denoted by μ, so that

$$\mu = \frac{\sin i}{\sin r}.$$

If it is desired to show that the light passes from air to glass, then the index of refraction is denoted by the symbol ${}_a\mu_g$. In general, if the light traverses from one medium (1) to a second (2), the index of refraction of the second medium with respect to the first is denoted by ${}_1\mu_2$. For most purposes we may assume $\mu = {}_a\mu_g$, etc.

If μ is the absolute refractive index of a medium, the quantity $(\mu - 1)$ is known as its ***refractivity.***

Experimental Determination of μ for a Plate of Glass.—Let AB and XY, Fig. 20·1, be the two parallel faces of a block of glass, and let two pins [shown by small black circles] indicate the incident ray OC. The position of the ray after passing through the glass is found by looking along the direction KD, which is marked by two more pins, the four pins being placed so that they are apparently collinear. If the block is now removed and the points C and D joined together by means of a straight line, the path of the ray of light is completely defined by OCDK. The angles of incidence and refraction at C are indicated by i and r respectively. By measuring these angles and using trigonometrical tables, μ can be calculated. If DE is normal to the face XY, the $\widehat{EDC}$ is also equal to r. Let e be the angle of emergence at D. Then

$$\frac{\sin i}{\sin r} = {}_a\mu_g \quad . \quad . \quad . \quad (1)$$

But the ray of light KD would traverse the medium along the path KDCO—a fact which is easily verified by looking along OC—so that

$$\frac{\sin e}{\sin r} = {}_a\mu_g \quad . \quad . \quad . \quad (2)$$

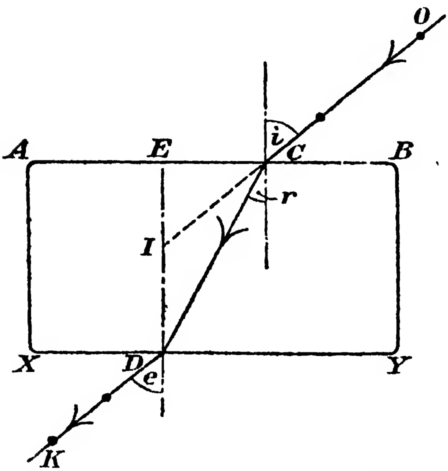


FIG. 20·1.—Path of a Ray of Light through a Parallel Plate.

The two fractions (1) and (2) are equal, so that

$$e = i.$$

A Second Method of Calculating the Value of μ.—Instead of measuring the angles i and r and determining μ from the ratio of the sines of these angles, it is better to produce the ray OC to cut DE in I. Then $\widehat{EIC} = i$, so that

$${}_a\mu_g = \frac{\sin i}{\sin r} = \frac{\sin \widehat{EIC}}{\sin \widehat{EDC}} = \frac{\dfrac{EC}{CI}}{\dfrac{EC}{CD}} = \frac{CD}{CI}.$$

Thus the ratio of the lengths of CD and CI is ${}_a\mu_g$; and it is much more easy and convenient to measure lengths than it is to measure angles.

Geometrical Construction for the Refracted Ray.—When the angle of incidence is given it is easy to determine the refracted ray, if μ is known. For example, let AO, Fig. 20·2, be a ray in air incident at O upon a plane surface of separation M_1M_2. With O as centre and any radius describe an arc of a circle EF to cut the ray in G; with O as centre and radius OC, where $OC = \mu OE$, describe a second arc CD. Through G draw GN perpendicular to

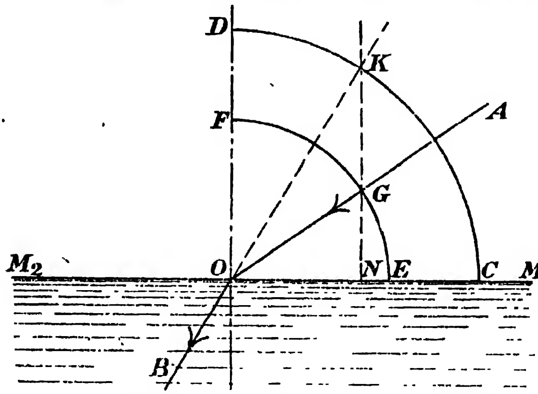


FIG. 20·2.—Geometrical Construction for a Refracted Ray.

the surface to cut CD in K. Produce KO to B, then OB is the refracted ray.

$$\text{Now} \qquad \mu = \frac{\sin i}{\sin r} = \frac{\sin \widehat{OGN}}{\sin r} \; [\because \widehat{OGN} = \widehat{AOD} = i]$$

where r is the angle of refraction which has to be discovered. Hence

$$\sin r = \frac{\sin \widehat{OGN}}{\mu} = \frac{\frac{ON}{OG}}{\mu} = \frac{ON}{OK} \; [\because \mu OG = \text{radius OK}].$$

$\therefore$ r is the $\widehat{OKN}$,

$\therefore$ Refracted ray is KO produced, i.e. OB.

Refraction through Several Media having Parallel Interfaces.—Suppose that the two media represented in Fig. 20·3 are water and glass; further, let air be the common medium which is all round these other media. Then denoting the angles as shown,

$$_a\mu_w = \frac{\sin i_1}{\sin r_1}$$

$$_w\mu_g = \frac{\sin r_1}{\sin r_2}$$

$$_g\mu_a = \frac{\sin r_2}{\sin r_3},$$

i.e. $$_a\mu_w \cdot {}_w\mu_g \cdot {}_g\mu_a = 1,$$

because it is an experimental fact that $i_1 = r_3$ when the interfaces are parallel.

But
$$_g\mu_a = \frac{1}{_a\mu_g},$$
hence
$$_w\mu_g = \frac{1}{_a\mu_w \times {}_g\mu_a} = \frac{_a\mu_g}{_a\mu_w},$$
i.e. the index of refraction of a third medium with respect to the second is equal to the refractive index of the third with respect to air, divided by the refractive index of the second with respect to air. For the refraction occurring at the water-glass interface we have

$$_w\mu_g = \frac{\sin r_1}{\sin r_2}.$$

Since $_w\mu_g = \frac{_a\mu_g}{_a\mu_w}$, the above equation becomes

$$_a\mu_w \sin r_1 = {}_a\mu_g \sin r_2.$$

This equation assumes a form which is more easily remembered, if we call $_a\mu_w = \mu_1$, and $_a\mu_g = \mu_2$. Further, let r_1 and r_2 be replaced by θ_1 and θ_2 respectively; then the equation becomes

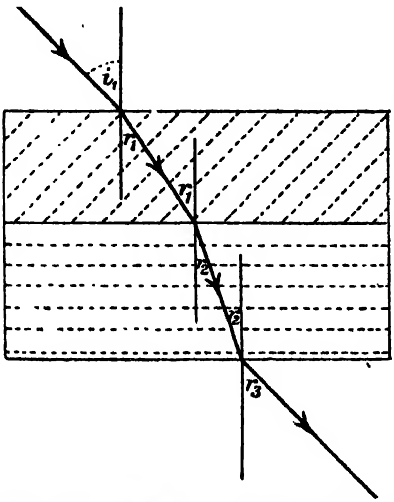


FIG. 20·3.—Refraction through Several Media.

$$\mu_1 \sin \theta_1 = \mu_2 \sin \theta_2.$$

It shows at once that when a refraction of light occurs the quantity $\mu \sin \theta$ remains constant. It is called an ***optical invariant.***

The above equation may be regarded as the fundamental equation of geometrical optics: it expresses the second law of refraction in its most general form. Moreover, if we write $\mu_1 = -\mu_2$ (this must be regarded as a mathematical operation) we have $\sin \theta_1 = -\sin \theta_2$, i.e. $\theta_1 = -\theta_2$, which is the second law of reflexion. For if AO, Fig. 20·4 is a ray of light, O a point on a reflecting surface, and if ON is the normal incident at that

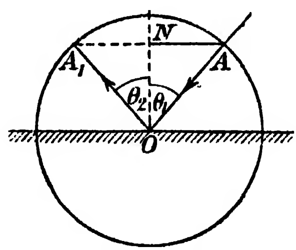


FIG. 20·4.—The Second Law of Reflexion.

point, we have, in the usual way, $\sin\theta_1 = \dfrac{NA}{OA}$. Then

$$\sin\theta_2 = -\frac{NA}{OA} = \frac{NA_1}{OA_1},$$

where $|NA_1| = NA$, so that θ_2 is the angle shown in the diagram.

Image formed by Refraction at a Plane Surface.— Let O, Fig. 20·5 (*a*), be a small object in any medium, and ON the normal through O to the surface of separation of the two media. It is desired to determine the position of the image of O as seen by an eye directed along NO. Suppose that the lower medium is glass and that the upper one is air. The ray OA,

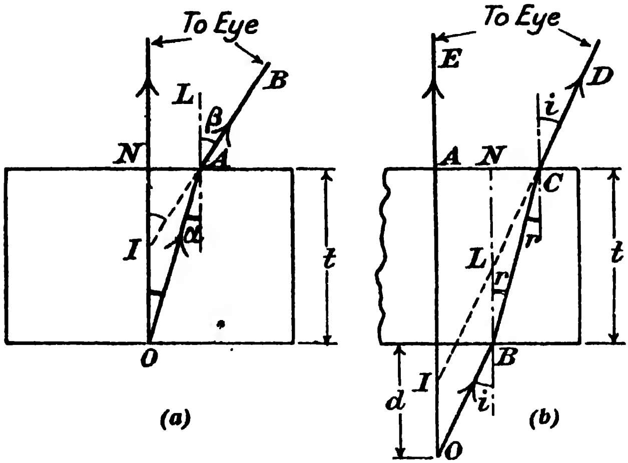


FIG. 20·5.—Formation of Image by Refraction.

incident on the upper face at an angle α, is refracted away from the normal at A, i.e. the refracted ray is AB. Let BA produced meet ON in I. The normals to the interface at N and A being parallel,

$$_g\mu_a = \frac{\sin\alpha}{\sin\beta} = \frac{\sin AON}{\sin AIN} = \frac{IA}{OA} \text{ [as on p. 394].}$$

Now, since the pupil of the eye is small, it follows that if the ray AB is to enter the eye, then the ray OA must be very nearly parallel and equal to ON, whilst IA is nearly equal to IN. Since OIN is also a ray of light from O, the image must be at I. Under these conditions

$$_g\mu_a = \frac{\sin\alpha}{\sin\beta} = \frac{IN}{ON},$$

or $$_a\mu_g = \frac{ON}{IN} = \frac{\text{actual thickness of block}}{\text{apparent thickness}} = \mu \text{ (say).}$$

The amount by which the object appears to be displaced from its true position is

$$OI = (ON - NI) = ON\left[1 - \frac{NI}{NO}\right]$$
$$= ON\left[1 - \frac{1}{\mu}\right].$$

To Determine the Refractive Index of a Liquid available in Large Quantities.—Suppose that water is the liquid. A tall glass vessel, A, Fig. 20·6, is filled with water. B is a cylindrical metal box turned upside down and having a long slot in its base. This slot is horizontal and illuminated with the aid of an electric lamp. Vertically above the slot and parallel to it is a nickel wire, C, heated to redness by an electric current. The position of the wire is adjusted until its image formed by reflexion in the surface of the water coincides with the image of the slot formed by refraction at that same surface. If d_1 and d_2 are the distances indicated, d_2 is the apparent depth of the illuminated slot when viewed directly from above. Hence

$$\mu = \frac{d_1}{d_2}.$$

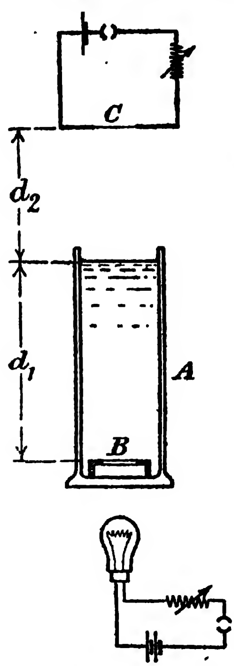


FIG. 20·6.—Refractive Index of a Liquid (available in bulk) by the Method of Apparent Depth.

The experiment succeeds more readily if the lamp is screened and the intensity of the light reduced somewhat with the aid of a resistance placed in the battery circuit.

Microscope Method for the Determination of μ.—A microscope whose objective has a working distance of one to three inches is required [this means that when the objective of the microscope is at this distance from an object a clear image is seen]. It should be capable of movement along a vertical scale attached to a stand and its position should be given by a vernier. Make a pencil-mark on a piece of paper and stick it to the bench ; focus the microscope, with its axis vertical, on this mark and read the vernier. Put a thick block of glass on the paper, e.g. a cubical paper weight ; as the mark is now apparently raised a distance OI, Fig. 20·5 (*a*), it is no longer in focus. Move the microscope along the scale until the mark is clearly seen, and again read the vernier. Finally scatter a

few grains of chalk on the upper surface of the block; focus the microscope on these and again observe its position. The instrument has now been focused in succession on points corresponding to O, I, N, Fig. 20·5 (*a*), hence the difference between the first and last readings gives the distance ON, and that between the second and third IN; ${}_a\mu_g$ can therefore be found. The same method may be applied to find the refractive index of a liquid. A piece of lead with a scratch on it is placed in a beaker and the microscope focused on the scratch as before. Liquid is then poured in, care being taken not to move the lead, and the microscope focused in succession on the mark and on chalk grains floating on the liquid surface. The calculation is made as in the last example.

Determination of the Displacement due to Viewing an Object normally through a Glass Plate.—Let O, Fig. 20·5 (*b*), be a small object at distance d from the nearer surface of a glass parallelopiped. Suppose that E is an eye viewing this object along a normal OAE. The size of the eye is very much exaggerated in the diagram to enable a clear figure to be constructed. Let OBCD be a ray of light passing through the glass. As on p. 397, the position of the image may be found by producing DC to meet OA in I. Since OA is also a ray from O it follows that I must be the image. To calculate the magnitude of the shift OI we note that CI meets BN, the normal at B, in L. Now L would be the image of an object B, and we have already seen [cf. p. 398] that $BL = t\left[1 - \frac{1}{\mu}\right]$, where t is the thickness of the glass. But OBLI is a parallelogram, so that $BL = OI$. The shift produced is therefore $t\left[1 - \frac{1}{\mu}\right]$, which is independent of the distance of the object from the glass block.

Determination of the Index of Refraction of the Material of a Thick Mirror.—Let MM, Fig. 20·7 (*a*), be the silvered surface of a glass mirror of thickness t, and index of refraction μ. Let O be an object [a vertical pin] placed on ON the normal to the mirror through O. If OA is a ray of light very close to ON, the refracted ray AB will be reflected from the back surface as the ray BC which emerges from the block as the ray CD. Moreover, the ray ON will be reflected from the surface MM and return along NO. If the rays CD and NO enter an observer's eye an image of the pin will appear at I. This image may be located by placing a second pin behind the mirror in such a position that it appears to coincide with I even when the eye is displaced from side to side. There is then said to be no *parallax* between this pin and the image I. Let α and β be the angles indicated, while u and v are respectively the distance of the object in front of and the distance

of the image behind the unsilvered surface of the mirror. Then

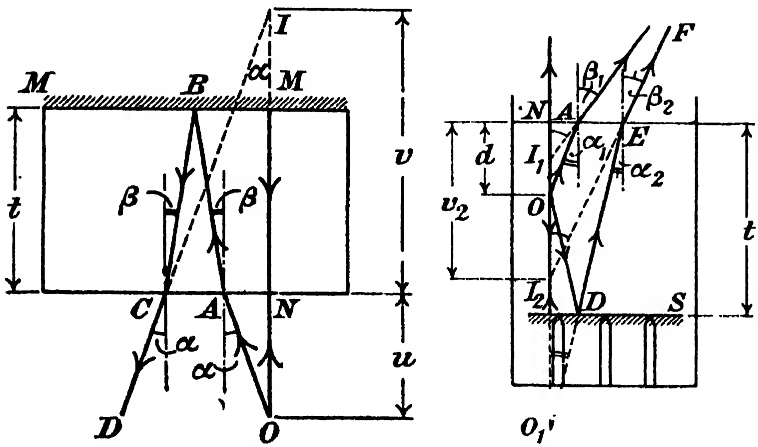
$$CN = CA + AN,$$

i.e. $$v \tan \alpha = 2t \tan \beta + u \tan \alpha.$$

Since α and β are small, $\tan \alpha \div \tan \beta = \mu$, so that

$$\mu v = 2t + \mu u.$$

Hence when u, v, and t are known μ may be calculated.



(a) (b)

FIG. 20·7.

Example.—Suppose that O, Fig. 20·7 (*b*), is a luminous point at distance d below the surface of a liquid whose refractive index is μ. A silvered surface, S, is placed in a horizontal position at a depth t in the liquid. It is required to find the positions of the images seen by an observer looking along the normal NO. The first image is formed by rays of light such as OA which are refracted at the surface of the liquid and appear to come from I_1. We have already determined the position of this image [cf. p. 397].

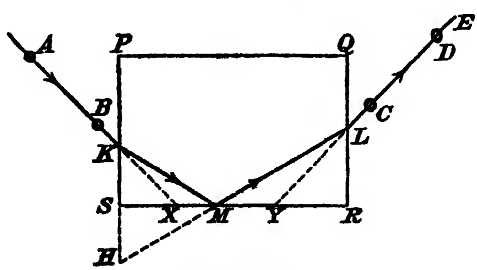


FIG. 20·8.

The second image I_2 is formed by rays of light such as OD which are reflected from the mirror along DE and finally emerge after refraction at the free surface of the liquid along EF. To calculate the position of this image we note that the ray DE apparently proceeds from O_1, the image of O in the silvered surface. Hence $\mu NI_2 = NO_1$,

i.e. $$v_2 = NI_2 = \frac{NO_1}{\mu} = \frac{2t - d}{\mu}.$$

Experiment.—Two pins, A and B, Fig. 20·8, placed as shown with reference to a rectangular block of glass PQRS, are viewed by an eye at E. Pins C and D are used to indicate the path of the ray after refraction at K, internal reflexion at M (but not necessarily *at* the critical angle, cf. below), and refraction at L. To show the path of the light ray through the glass, an outline of the block having been drawn, AB and CD are produced to cut this outline at K and L, and again at X and Y. To determine M, PS is produced so that KS = SH. M is obtained by joining HL. The refractive index of the glass is equal to the ratio $\frac{KM}{KX}$, or to $\frac{LM}{LY}$ [cf. p. 394].

Total Internal Reflexion.—Let us now consider what happens when light passes from a dense to a rare medium. Let O, Fig. 20·9,

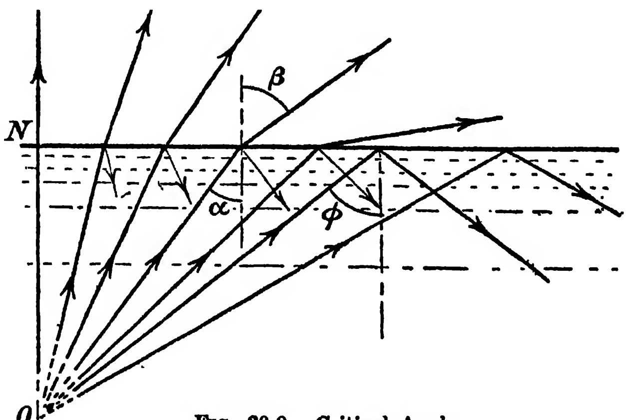


FIG. 20·9.—Critical Angle.

be a luminous point in a medium whose absolute refractive index is greater than unity (water, for example), ON being the normal through O to the surface. If α and β are the angles indicated, then, with the usual notation,

$$_w\mu_a = \frac{\sin \alpha}{\sin \beta} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

or

$$_a\mu_w = \frac{\sin \beta}{\sin \alpha} = \mu \text{ (say)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This latter fraction is greater than unity. Since $\sin 90° = 1$, it follows that the emergent ray will travel along the surface when

$$\frac{1}{\sin \alpha} = \mu, \text{ or } \sin \alpha = \frac{1}{\mu} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The rays of light which travel from O and which are incident upon the surface at a greater angle of incidence than that given by (3)

cannot pass into the air—they are ***totally internally reflected*** and obey the usual laws of reflexion. The particular angle of incidence given by (3) is termed the ***critical angle*** for rays travelling in water and incident upon a water-air interface. If it is denoted by ϕ,

$$\mu \sin \phi = 1 \text{ , or cosec } \phi = \mu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

In the general case, the term 'critical angle' refers to rays travelling in a medium whose refractive index ***relative*** to that of a second medium is greater than unity and incident upon an interface between the two media. It will be noted that total internal reflexion is only possible on the side of an interface where the medium has the higher absolute refractive index.

It must be realized quite clearly that although it is only when the angle of incidence is greater than the critical angle that the light is totally internally reflected, yet, for angles of incidence less than the critical angle, a portion of the light is reflected, the remainder being refracted—cf. the diagram.

Experimental Determination of the Refractive Index of Glass.—Suppose that A, Fig. 20·10 (a), is a semi-cylindrical block of glass, and that B is a straight black line ruled on a piece of ground glass, C. Suppose that B is ***almost*** in contact with A along its axis. By viewing the line B through the curved surface

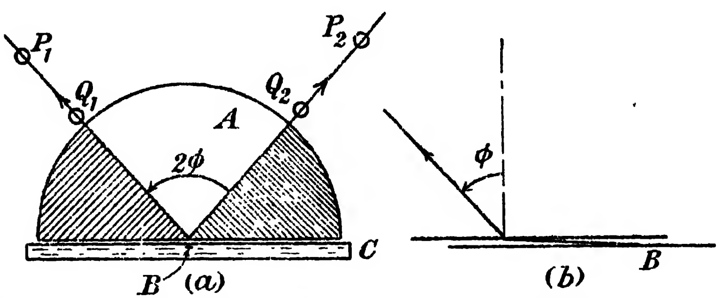


FIG. 20·10.—A Critical Angle Method for Determining the Refractive Index of Glass, ${}_a\mu_g$.

of the glass the directions of a series of rays emerging from the curved surface of A may be traced. It will be found that the field of view has two definite limits, beyond which it is impossible to see the line B. Let pins P_1Q_1 and P_2Q_2 define these limits. If these rays are produced backwards they will intersect at B, the angle between them being 2ϕ, where ϕ is the critical angle for rays incident upon a glass-air interface.

To discover the reason for this, let us consider Fig. 20·10 (b).

This is a very much enlarged diagram of the region near B and shows a ray of light from B striking the flat edge of A at grazing incidence. The refracted ray then makes an angle almost equal to ϕ with the normal at the point of incidence, and when the grazing angle is zero, the angle of refraction is ϕ. Then

$$\operatorname{cosec} \phi = {}_a\mu_g.$$

To Determine the Refractive Index of Water by a Critical Angle Method.—Let ABCD, Fig. 20·11, be a ray of light passing through a rectangular block of glass bounded on two sides by water and by air respectively. Let α, β, and γ be the angles indicated. Then ${}_a\mu_w \sin \alpha = {}_a\mu_g \sin \beta$, and ${}_a\mu_g \sin \beta = \sin \gamma$, i.e. ${}_a\mu_w \sin \alpha = \sin \gamma$. As α increases, γ finally reaches a value $\frac{\pi}{2}$, when the ray BC is internally reflected at C. If ϕ is the value of α when this occurs, ${}_a\mu_w \sin \phi = 1$, i.e. ϕ is the critical angle for water-air. Hence ${}_a\mu_w$ may be calculated when ϕ is known. The necessary apparatus is indicated in Fig. 20·12. Two small plates of glass, A, are cemented together by sealing-wax along their edges so that an air film of constant thickness remains between the plates. This is attached to a pointer moving over a circular scale [shown dotted]. A slit S, illuminated by a sodium flame, a converging lens arranged so that S is in a focal plane so that parallel rays pass through it, the compound plate A immersed in water, and a telescope T, focussed for parallel light, are arranged in a straight line. A is

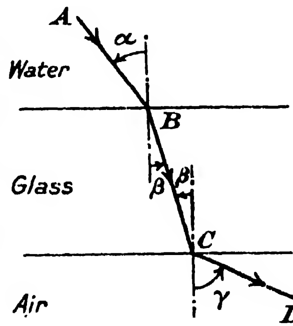


FIG. 20·11.

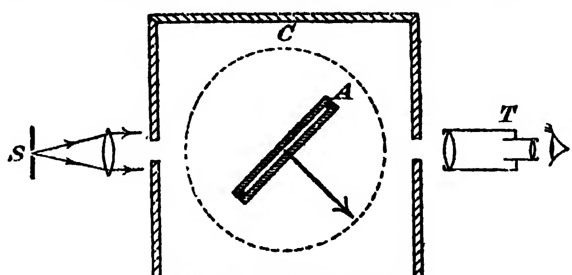


FIG. 20·12.—μ by a Critical Angle Method.

rotated until the image of S in the telescope just disappears. The position of the pointer having been noted, A is rotated until the

image appears again. The rotation is continued until darkness occurs again, and so on until A has moved through 360°. The positions of the pointer having been noted on each occasion, the mean value of the angle through which the plate may be moved and the field remain ***bright*** is deduced. Half this angle is ϕ, and we have already shown that ${}_a\mu_w \sin \phi = 1$.

[If, when total internal reflexion has occurred, the space between the plates is filled with water, an image of S appears at once in the telescope.]

The Refractive Index of a Liquid.—Fig. 20·13 shows a box to which two equal, upright, brass strips PN′ and QN are fixed; a scale in mm. forms the base for these uprights. The liquid is placed in the box and the screw A is moved until the surface of the liquid is just above N and N′—the scale is then level. Observe, through

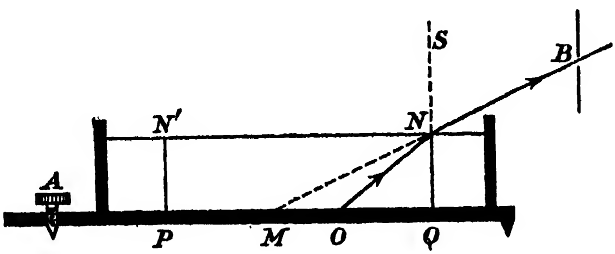


FIG. 20·13.—Measurement of μ for a Liquid.

a cardboard slit B, the division on the scale which is just visible—ONB is the path of the ray. Run off the liquid and again observe through B the division M, which can just be seen—BNM is a straight line. Measure QN and make a large scale drawing of QOMN. Now μ, the index of refraction of the liquid with respect to air, is expressed by

$$\mu = \frac{\sin \widehat{MNQ}}{\sin \widehat{ONQ}}.$$

If, therefore, these two angles are measured, the value of the refractive index can be calculated.

The refractive index of a liquid may also be determined with the aid of a concave mirror. Its radius of curvature is first measured by making use of the fact that if a small object such as a pin is at the centre of curvature of a concave mirror its image is also there. A small quantity of liquid is then introduced into the mirror, and a pin placed in a horizontal position is moved up

and down until its point is again at the same distance from the mirror as its image. In making this adjustment it is best to work with the point of the pin owing to distortion produced near the other end of the image. This distortion is caused by the curved surface of the liquid near the periphery of the mirror. Let C and O, Fig. 20·14, be the positions of the pin in the two instances respectively. Then C is the centre of curvature of the mirror. The image O is produced by rays such as OA which after refraction at the surface of the liquid travel along AB, a normal to the surface of the mirror. Such rays are reflected along their original paths and form an image at O. If i and r are the angles of incidence and refraction at A,

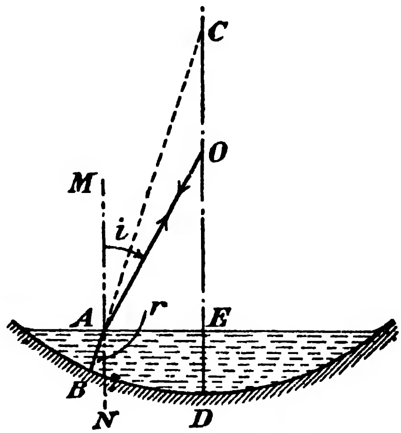


FIG. 20·14.—Refractive index of a Liquid by means of a Concave Mirror.

$$\mu = \frac{\sin i}{\sin r} = \frac{AE}{AO} \div \frac{AE}{AC} = \frac{AC}{AO}.$$

If the mirror has a large radius of curvature the depth of the liquid is small and we may assume that OA = OE = OD, and

$$AC = EC = DC. \quad \text{Hence } \mu = \frac{CD}{OD}.$$

The Principle of a Pulfrich Refractometer.—The Pulfrich refractometer is an instrument used to determine the refractive indices of liquids, oils, fats, and, with a slight modification, transparent solids. Its essential optical part consists of a glass cube G, Fig. 20·15 (a), of known refractive index μ_2. The upper portion of this cube has been ground away in part so that a truncated cone is left—cf. Fig. 20·15 (b). AB is a shallow glass vessel cemented to the upper face of the worked cube: it contains the liquid whose refractive index μ_1 is to be determined. Monochromatic light [cf. p. 449] from a source S is directed by means of a converging lens L to be incident along the upper face of the original cube. Consider a ray which enters the cube at C so that the angle of refraction is ϕ, the critical angle for rays of light travelling in glass and incident upon the glass-liquid interface. If CE is the path of this ray inside the cube, let the corresponding emergent ray be EK,

the angle of emergence being θ. Suppose that this angle can be measured. To establish the equation from which the refractive index of the liquid may be calculated, the second law of refraction in its general form $\mu_1 \sin \theta_1 = \mu_2 \sin \theta_2$ [cf. p. 396] will be used. The light rays, which are incident on the glass-liquid interface at an angle just less than 90°, are refracted into the glass at the critical angle ϕ, so that they travel in the direction CE, and after refraction at E emerge along EK.

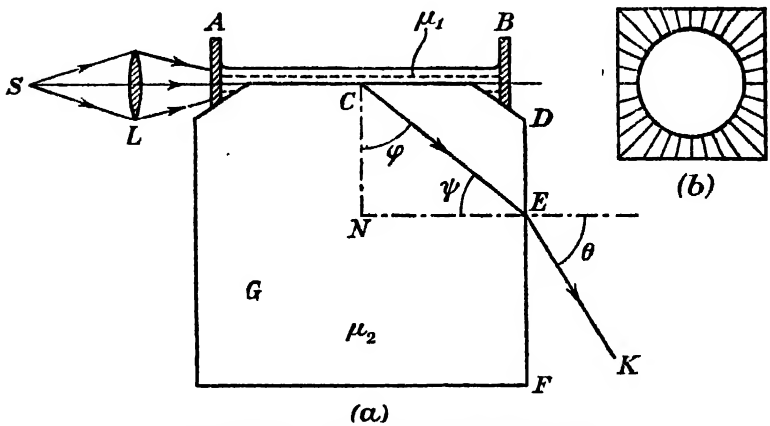


FIG. 20·15.—The Pulfrich Refractometer.

Since μ_1 and μ_2 are the absolute refractive indices of the liquid and glass respectively, for the refraction of light at C, $\theta_1 = \frac{\pi}{2}$, and $\theta_2 = \phi$, where θ is the critical angle for rays travelling in glass and incident upon a glass-liquid interface, so that $\mu_1 \sin \frac{\pi}{2} = \mu_2 \sin \phi$,

i.e. $$\mu_1 = \mu_2 \sin \phi \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

The angle of incidence corresponding to the angle of emergence θ is

$$\widehat{CEN} = \psi = \left(\frac{\pi}{2} - \phi\right).$$

The general law of refraction gives in this instance,

$$\mu_2 . \sin\left(\frac{\pi}{2} - \phi\right) = 1 . \sin \theta,$$

or $$\mu_2 . \cos \phi = \sin \theta. \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

Squaring and adding the expressions for $\sin \phi$ and $\cos \phi$ in

order to eliminate ϕ from equations (i) and (ii) we get, since $\sin^2\phi + \cos^2\phi = 1$,

$$\left(\frac{\mu_1}{\mu_2}\right)^2 + \left(\frac{\sin\theta}{\mu_2}\right)^2 = 1,$$

or

$$\mu_1{}^2 = \mu_2{}^2 - \sin^2\theta. \quad . \quad . \quad . \quad \text{(iii)}$$

Equation (iii) shows that for ϕ to be real, μ_1 must be less than μ_2, i.e. the refractive index of the liquid must be less than that of the glass forming the cube. Further, since the extreme limits for ϕ are $0 \rightarrow \frac{\pi}{2}$, the liquid under investigation must have a refractive index within the range $\sqrt{\mu_2{}^2 - 1} \rightarrow \mu_2$.

Caustic Curve by Refraction at a Plane Surface.—Let XY, Fig. 20·16, be the trace of a plane interface between two media, the lower one having a refractive index μ with respect to the upper one.

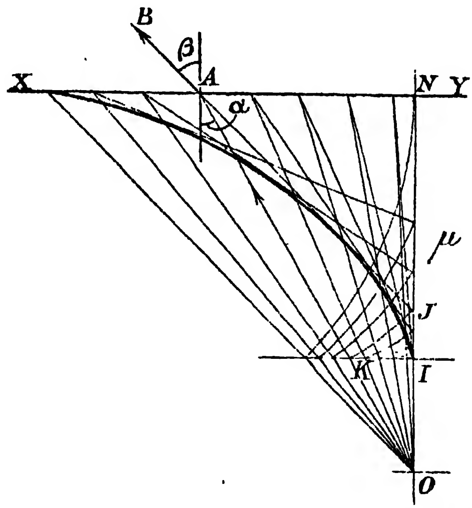


FIG. 20·16.—Caustic Curve by Refraction at a Plane Surface.

Any ray of light such as OA emitted from the luminous point O. is refracted at the interface and then travels along the direction AB, Suppose that BA produced cuts the normal NO in J. As A moves away from N the point J moves towards N. If many such paths are constructed for various positions of A, then it is found that AJ is always tangential to a certain curve termed the caustic curve by refraction at a plane surface. This curve is the envelope of all the lines AJ.

This particular caustic is very easily constructed. We take a point I in ON such that ON $= \mu$. IN, and through I draw a straight line IK parallel to XY. Draw any ray OA proceeding from the luminous point O, cutting XY in A and IK in K. With centre A and radius AK describe an arc to cut ON in J. Join JA and produce it to B; then AB is the refracted ray, for

$$\mu = \frac{ON}{IN} = \frac{OA}{AK} = \frac{OA}{AJ} = \frac{OA}{AN} \cdot \frac{AN}{AJ}$$
$$= \frac{AN}{AJ} \Big/ \frac{AN}{OA} = \frac{\sin\beta}{\sin\alpha},$$

where α and β are the angles of incidence and of refraction at A.

If a sufficient number of lines AJ are constructed their envelope is the caustic curve required.

Refraction through a Prism.—In optics the term ***prism*** denotes a transparent body bounded by plane polished surfaces which intersect in parallel straight lines. In a refracting prism only two plane surfaces are essential: these are termed its refracting surfaces. The light enters at the first surface and emerges at the second. The angle between the refracting surfaces is the ***refracting angle*** and their line of intersection the ***refracting edge.*** A section

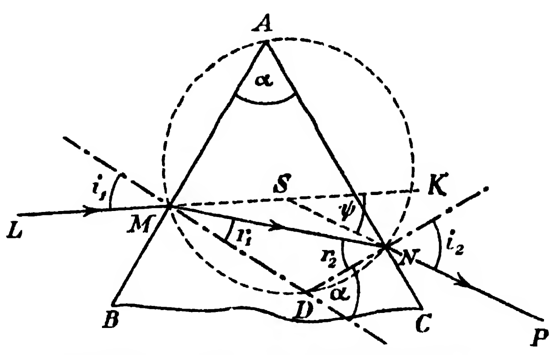


FIG. 20·17.—Refraction through a Prism.

of the prism made by any plane normal to one (and hence three) of its edges is termed a ***principal plane*** of the prism. Let ABC, Fig. 20·17, be a principal section through a prism—the base is drawn as shown in order to indicate that it plays no part in the present problem. The $\widehat{BAC}$ is called the angle of the prism (α). Let LM be the incident ray, i_1 being the angle of incidence. Let MN be the path of the ray in the prism, NP the emergent ray, and i_2 the angle of emergence. Produce PN to meet LM produced in S; in order to bring LS into the same direction as NP, it must

be rotated through the $\widehat{KSP}$,—this is called the angle of deviation and is denoted by the letter ψ.

Suppose that MD and ND are the normals to the faces of the prism at M and N; then a circle can be drawn to pass through A, M, D and N. It therefore follows that $\widehat{MDN} = 180° - \alpha$.

Also, since the exterior angle of a triangle is equal to the sum of the two interior and opposite angles,

$$\alpha = r_1 + r_2,$$

$$\begin{aligned}\psi &= \widehat{SMN} + \widehat{SNM} \\ &= (i_1 - r_1) + (i_2 - r_2) = (i_1 + i_2 - \alpha).\end{aligned}$$

Experiment.—Commencing with angles of incidence not less than 35°, and increasing them by 5° intervals to about 70°, plot the paths of light rays through a prism. Measure the angles of incidence and the corresponding angles of deviation. Two results are obtained from each setting since if i_2 is the angle of incidence, the deviation is still equal to ψ.

It will be found that ψ decreases and then increases, as i increases, i.e. there is a minimum value for ψ. This angle is called the ***angle of minimum deviation*** and will be denoted by the letter γ, i.e. $\psi_{min} = \gamma$. It will also be discovered that the deviation is a minimum when i_1, i_2, i.e. the angles of incidence and emergence, are equal, i.e. the ray passes symmetrically through the prism.

Image Produced by a Prism.—Let P, Fig. 20·18, be a luminous point and suppose that PQ is that ray which, after refraction, passes through the prism with minimum deviation. If PR and PS are two other rays incident at slightly different angles, an inspection of the graph obtained above shows that the deviation of these rays will

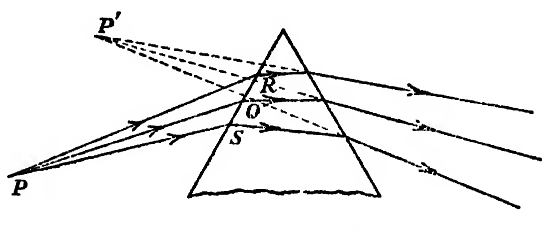


FIG. 20·18.—Image produced by a Prism.

be practically the same as that of PQ since near the minimum on the curve the deviation only varies slightly with the angle of incidence. Hence the passage of the rays through the prism does not alter the amount by which they diverge, i.e. the emergent rays appear to come from a common point P′. If the thickness of the prism is negligible compared with the distance of P from it, the

points P and P′ are equally distant from the prism. P′ is the virtual image of P. When the prism is not in the position of minimum deviation the rays emerging from it no longer intersect in a common point, i.e. no true image is formed.

Experimentally the position of P′ may be found by using a pin as object and placing a second pin so that there is no parallax between it and P′ [this second pin must be sufficiently long to be seen over the top of the prism].

Measurement of the Angle of a Prism.—Suppose that AB and AC, Fig. 20·19, are the two faces of a prism between which it is desired to measure the angle. Parallel straight lines having been ruled upon a sheet of paper, pins are placed at D, E, F and G to define two parallel rays. These rays are reflected from the prism faces

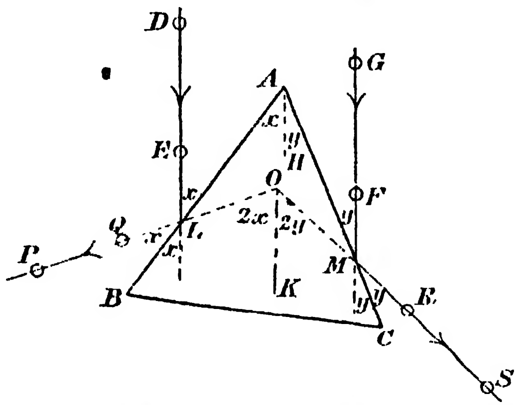


FIG. 20·19.—Determination of the Angle of a Prism.

and the reflected rays are defined by means of the pins P, Q, R and S. Let L and M be the points of incidence of the rays. Produce PQ and SR to meet in O, and then draw AH and OK parallel to the incident rays DE and GF.

Now all the angles marked x are equal, so that the $\widehat{LOK}$ is $2x$, because OK is parallel to DE produced. Similarly the $\widehat{MOK}$ is $2y$ where y is the angle indicated. It therefore follows, by simple addition, that the $\widehat{LOM}$ is twice the angle of the prism, for this latter is $x + y$.

Determination of the Refractive Index of the Material of a Prism.—***Method i*:** When the angle of incidence i corresponding to the minimum deviation γ has been found, μ can be calculated. The experiment is described on p. 455. For, in this instance, $i_1 = i_2 = i$, $r_1 = r_2 = r$ (say).

Hence $$\alpha = 2r \text{ or } r = \frac{\alpha}{2},$$

and $$\gamma = 2(i - r),$$

or $$2i = \gamma + \alpha, \text{ i.e. } i = \frac{\alpha + \gamma}{2}.$$

$$\therefore \mu = \frac{\sin i}{\sin r} = \frac{\sin \frac{1}{2}(\alpha + \gamma)}{\sin \frac{1}{2}\alpha}.$$

This equation involves α, the angle of the prism, so that this quantity must be known before the refractive index can be calculated.

Method ii: A piece of ground glass is attached by an india-rubber band to the base BC of a prism ABC, Fig. 20·20. Then every point on the base is a source of light sending rays in all directions. Let S be such a point and consider two rays SH and SK emitted by S in the principal plane of the prism. If SH is incident upon the face AB at an angle less than ϕ, the appropriate critical angle, there will be a refracted ray HJ. On the other hand, if SK is incident at the critical angle ϕ, it will be totally reflected from the face AB and strike the face AC from which it will emerge in the direction PQ. Similarly, rays from S incident upon AB at angles greater than ϕ will leave the prism after being refracted at the face AC. If, therefore, an observer looks into the face AC he will find that the field of view is divided into two regions, one relatively much darker than the other—it is the presence of extraneous light and that reflected from BA before the critical angle is reached (cf. p. 402) which prevent the field from being completely dark. The line of demarcation between the two regions will vary with the position of the observer, because different points in BC correspond to the particular line of demarcation observed. For one convenient position the line of demarcation may be indicated by two pins P and Q.

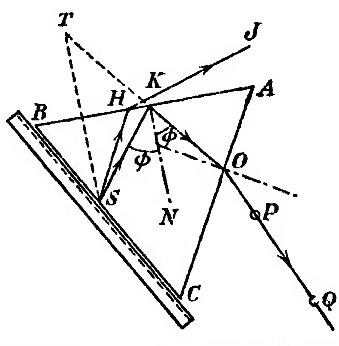


FIG. 20·20.—μ for Material of Prism by Critical Angle Method.

To determine the point in S from which the ray PQ proceeds a vertical line having been drawn upon the surface of the ground glass and this replaced, the glass is moved so that the line on it is always parallel to the refracting edge of the prism and until the image of the line formed by reflexion at K and refraction at the face AC appears to lie along QP produced. If T is the image of

S in AB, by joining TO by a straight line we obtain K. If KN is normal to AB at K, SKN is ϕ, the critical angle required, so that μ is known for $\mu \sin \phi = 1$.

A Critical Angle Method for Determining the Refractive Index of a Liquid.—The above experiment is readily modified so that the refractive index of a liquid may be determined. A piece of brass to fit over the side AB of the prism is first made—cf. Fig. 20·21. The central portion of this brass piece is removed to a depth of about 1 mm. A small quantity of the liquid under investigation is then placed in the hollows in the brass and the prism placed in contact with the liquid and its ' cell '. The line

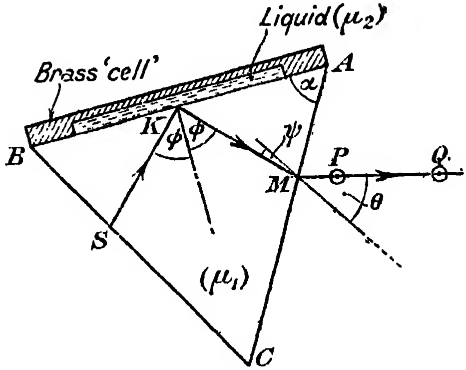


FIG. 20·21.—A Critical Angle Method for determining the Refractive Index of a Liquid.

of demarcation corresponding to a point S is then found as before in the previous experiment. If ϕ is the critical angle for rays of light travelling in glass and incident upon the glass-liquid interface, we have

$$\mu_1 \sin \phi = \mu_2 \sin \frac{\pi}{2},$$

[cf. p. 396], where μ_1 and μ_2 are the refractive indices of the glass and the liquid respectively. Hence if μ_1 is known and ϕ measured, μ_2 is easily calculated.

If, however, the angle α is known and the angle of emergence θ measured, μ_2 may be calculated from the formula now established. We have

$$\mu_1 \sin \phi = \mu_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

Let ψ be the angle of incidence at M. Then

$$\mu_1 \sin \psi = \sin \theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

Also, from the $\triangle$ AKM,

$$\alpha + \left(\frac{\pi}{2} + \psi\right) + \left(\frac{\pi}{2} - \phi\right) = \pi.$$

$$\therefore \phi = \alpha + \psi \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

From (i) we have,

$$\begin{aligned} \mu_2 &= \mu_1 \sin(\alpha + \psi) \\ &= \mu_1 \sin\alpha \cos\psi + \mu_1 \cos\alpha \sin\psi \\ &= \mu_1 \sin\alpha \sqrt{1 - \frac{\sin^2\theta}{\mu_1^2}} + \cos\alpha \sin\theta \\ &= \sin\alpha \sqrt{\mu_1^2 - \sin^2\theta} + \cos\alpha \sin\theta. \end{aligned}$$

Hence μ_2 may be calculated when μ_1, α, and ϕ are known. It should be noted that sometimes θ is on the side of the normal other than that indicated: then θ is negative, and we have

$$\mu_2 = \sin\alpha\sqrt{\mu_1^2 - \sin^2\theta} - \cos\alpha \sin\theta.$$

Prisms with Small Refracting Angles.—Suppose that α, the refracting angle of the prism, is small and that the incident and emergent rays are nearly normal to the respective surfaces. Then, as on p. 409, we have

$$\psi = (i_1 + i_2) - (r_1 + r_2),$$

and $r_1 + r_2 = \alpha$. Since i_1 and i_2 are small, r_1 and r_2 are also small, so that we may write instead of the general equation $\sin i = \mu \sin r$, $i = \mu r$, i.e. $i_1 = \mu r_1$ and $i_2 = \mu r_2$, and hence

$$\psi = (\mu - 1)(r_1 + r_2) = (\mu - 1)\alpha.$$

The above equation is used in connexion with achromatic prisms [cf. p. 465].

When the angle of incidence is large, however, the deviation produced by refraction through a prism whose refracting angle is small is not given by the above simple formula. For example, suppose $\alpha = 3°$, $i_1 = 60°$ and $\mu = 1{\cdot}523$. Then under such conditions $r_1 - r_2 = \alpha$ and $\psi = i_1 - i_2 - (r_1 - r_2) = i_1 - i_2 - \alpha$. Now calculation shows that $r_1 = 34°\ 40'$, $r_2 = 31°\ 40'$, so that $i_2 = 53°\ 4'$. Thus $\psi = 3°\ 56'$, i.e. the deviation is still small, but it is not equal to $(\mu - 1)\alpha$ or $1°\ 34'$, which is the deviation when the angle of incidence on the first surface of the above prism is small.

The Constant Deviation Spectrometer.—Let ABC, Fig. 20·22 (*a*), be one half of a 60° glass prism and suppose that PQ is a ray of light incident at such an angle, θ_1, that the refracted ray is parallel to BC. Let DEF be the other half of the prism placed as shown. The ray QR will emerge without change of direction. Assume it falls on a plane mirror M so placed that the angle of incidence is 45°. The reflected ray will enter the

prism DEF in a direction normal to BE and therefore parallel to EF. The emergent ray will be ST and the angle of emergence is θ_1. Thus the ray of light will pass through the compound arrangement of prism and mirror just as if it had passed through the original 60° prism in the position of minimum deviation. The difference,

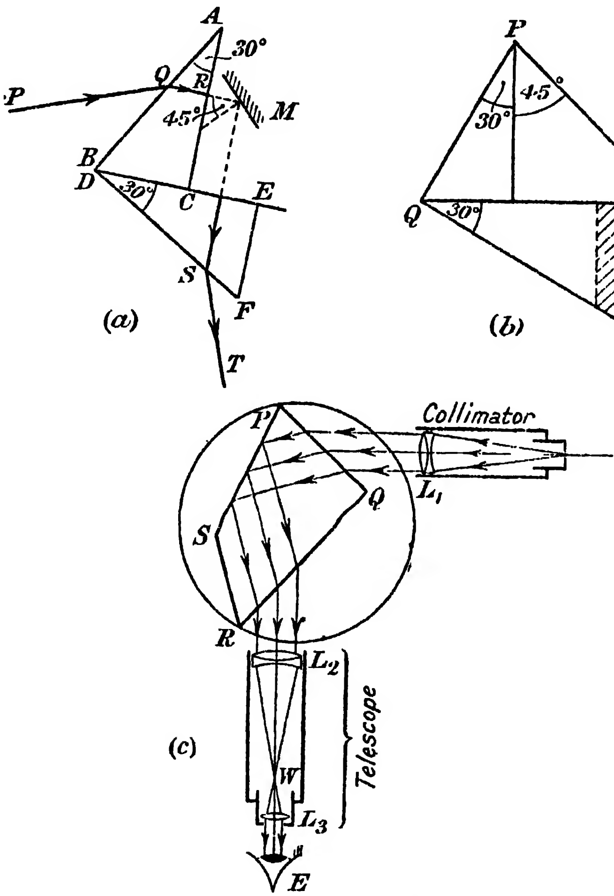


Fig. 20·22.—A Constant Deviation Spectrometer.

however, is that ST will be normal to PQ. [The proof is left as an exercise.]

If, instead of using a plane mirror, a 45° prism is placed as shown in Fig. 20·22 (*b*), the same effect will be produced : and the portion of glass shown shaded can be added so that the prism may be manufactured as an entity.

Suppose now that a collimator and telescope are fixed at right

angles to each other as shown in Fig. 20·22 (*c*). Then if monochromatic light is used the prism PQRS, constructed as above, can always be placed in such a position that the image (a spectral line) always falls on the cross-wires, W, of the telescope—the table on which the prism stands can be rotated about an axis parallel to the slit in order to bring the prism into the desired position. When heterogeneous light is used each line in the spectrum can be observed in turn on the cross-wires as the table is rotated. The rotation is brought about by means of a fine steel screw, whose point pushes against an arm projecting from the table, and whose head is in the form of a drum calibrated so that the wave-length of the line under observation may be read off directly. The scale is standardised by using lines of known wave-length: those in the mercury and copper area are very convenient. Of course an ordinary spectrometer may be calibrated but the translation of the telescope settings into wave-lengths is tedious. The constant deviation spectrometer avoids this time-consuming process, and should the prism be displaced it is easily replaced by using one or more known spectral lines—e.g. one of the two sodium D lines.

Images in a Thick Mirror.—To explain the formation of the several images seen when a candle is held in front of a thick mirror, let us consider what happens to a *single* ray OA, Fig. 20·23 (*a*), sent out by a luminous point O. At A, a point on the front surface of the mirror, a portion of the light energy is reflected giving rise to the ray AP, whilst a second portion is refracted giving the ray AB. At B, a point on the back of the mirror, reflexion occurs and we have the ray BC. When this reaches the front surface there is formed the reflected ray CD, parallel to AB, and the refracted ray CQ, parallel to AP. The further course taken by the light ray is indicated, and the diagram shows that there is a system of parallel rays emerging from the mirror. But this system does not produce the multiple images seen in a thick mirror. For an image to be seen there must be a pencil of light proceeding from the object to the eye of the observer. Let us consider a pencil of rays of which OA, Fig. 20·23 (*b*), is the central ray, and Oa_1 and Oa_2 the extreme rays. These rays will be partly reflected at the first surface, but we shall assume that these rays do not enter an eye E. Suppose, however, that the central ray is refracted along AB, reflected at B along BC, and that a refracted ray CD is produced at C. If the refracted pencil of which CD is the central ray is a_2 d_2 d_1 c_1 and this pencil enters the eye E, an image will be produced at I. This point is only *on* the normal to the mirror through O, if the eye is near to that normal. [In the diagram the distance between E and the above normal has been made large for the sake of clearness, although

actually the eye E is supposed to be near to the normal through O.]

To account for the multiple images let us refer to Fig. 20·23 (c) where only the central rays of different pencils from O have been drawn. If the pencil OA, after reflexion at A, enters an eye E an image will be seen at i_1. The rays belonging to this pencil which enter the glass traverse such paths that they do not enter the eye. Now the pencil OB likewise gives rise to a reflected pencil at the front surface, but these do not enter the pupil of the eye E. Let us asssume, however, that the rays emerging after one reflexion at the silvered surface do enter the eye: then a second image will be seen at I. This image will generally be the brightest, since most of the energy will be in the pencils which suffer one reflexion at the silvered surface. Similarly, if the rays in the pencil OC after two reflexions

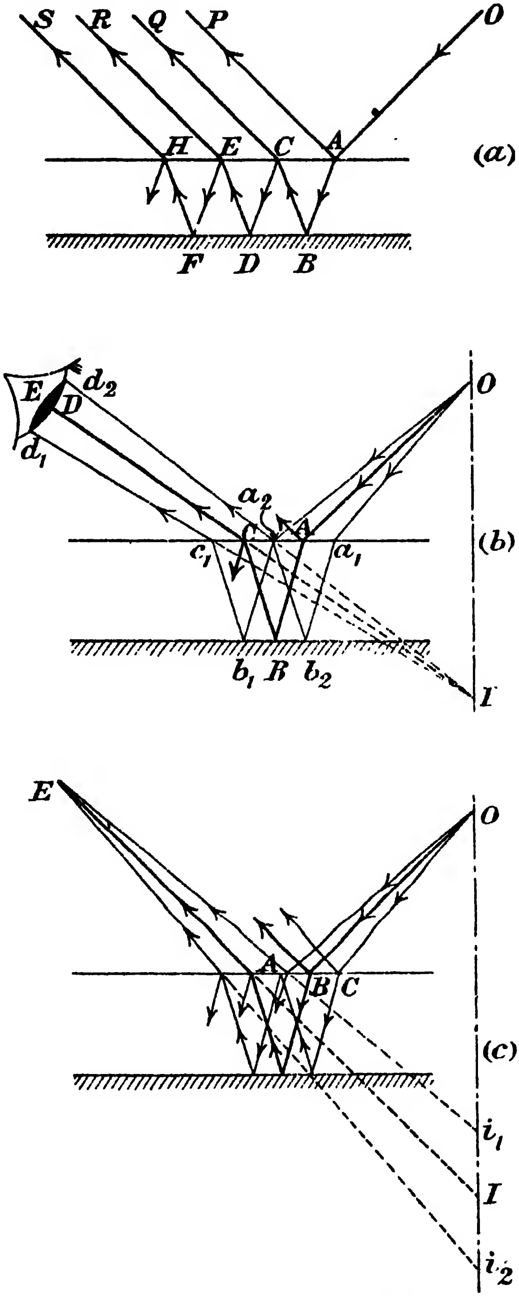


FIG. 20·23.—Images in a Thick Mirror.

at the back surface enter E a third image will be formed at i_3. The formation of the other multiple images may be explained in a similar manner. In practice, seldom more than six images are seen, for, owing to absorption in the glass, the energy in successive emergent pencils after the second rapidly diminishes.

It is interesting to note that when the angle of incidence increases, more and more energy is reflected from the first surface so that ultimately the first image becomes brightest.

An instructive variation of this experiment is to view the moon in a thick mirror, when only one image is observed. This is because all the pencils incident upon the mirror are parallel to one another, so that, in spite of the multiple refraction and reflexion of the rays, all the rays emerging from the mirror form a parallel system, and when rays belonging to such a system enter the eye only one image is seen. If a distant candle is observed in this way and several images are seen, the two faces of the mirror cannot be parallel to each other.

Atmospheric and Astronomical Refraction.—It sometimes happens that the layer of air immediately above a flat stretch of land or water on which the sun is shining is hotter than the more elevated layers. Its density and refractive index are therefore less. Rays of light from objects near the horizon are therefore incident on the surface at very large angles and are totally internally reflected. They therefore pass upward and may enter the eye of an observer, who then sees an inverted image of the object. This is a particularly annoying experience in a desert, for the image of the sky is often taken to be that of a stretch of water. Images produced in this manner are termed ***mirages***.

When distant objects are viewed through the hot air rising from a heated surface—a steam boiler, or a road on a hot day—they appear to move in a vibratory manner. Patches of hot air act like prisms deviating the rays passing through them. Since the size of such prisms is continually changing, the deviations are not constant and the image appears to be that of a vibratory object.

The twinkling or scintillating of the stars is similarly attributed to changing inequalities of the refractive index of portions of the atmosphere.

Another effect of atmospheric refraction is to make the stars appear higher than what they really are, each layer of air making a contribution to the total deviation, i.e. the refraction does not occur at one particular interface and the rays follow a curved path.

Another phenomenon attributable to atmospheric refraction is known as the ' horizontal moon '—we refer to the enlarged appearance of the moon when the latter is near to the horizon.

Worked examples:

(i) A ray of light is incident at normal incidence on the hypotenuse face of an isosceles right-angle prism whose material has a refractive index 1·414, the plane of incidence is normal to the refracting edges of the prism. Trace the path of the ray. Also indicate the paths of rays slightly inclined to the above ray.

The critical angle, ϕ, for a medium-air interface is such that

$$1{\cdot}414 \sin \phi = 1,$$

i.e. $$\phi = 45°$$

Hence the ray PQ, Fig. 20·24 (*a*), which enters the prism at Q, travels without deviation at Q as QR and then suffers total internal reflexion at R and again at S: it finally emerges along STU in a direction parallel to PQ.

If P_1Q, Fig. 20·24 (*b*), is another ray which makes a small angle α

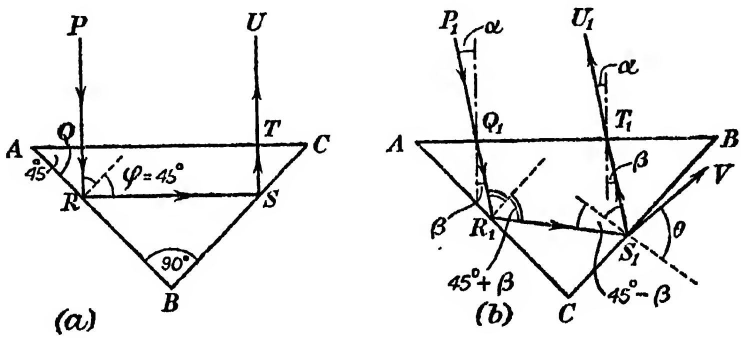


FIG. 20·24.

with the normal at Q_1, it will be refracted along Q_1R_1, where β, the angle of refraction, is determined by

$$\sin \alpha = \mu \sin \beta = 1{\cdot}414 \sin \beta.$$

Now Q_1R_1 is incident at an angle $45° + \beta$ at R_1, so that the ray is totally internally reflected as R_1S_1. At S_1 the angle of incidence is $45° - \beta$, which is less than the critical angle: thus there will be a refracted ray S_1V, the angle of emergence, θ, being determined by

$$1{\cdot}414 \sin (45° - \beta) = \sin \theta.$$

i.e. $\theta \longrightarrow 90°$.

The ray reflected at S_1 is S_1T_1 incident upon AB at an angle β, so that the refracted ray T_1U_1 is parallel to QP_1.

[Now try to trace the path of a ray P_2Q_1, where $P_1\widehat{Q_1}P_2 = 2\alpha$.]

(ii) If ϕ_1 and ϕ_2 are the angles of incidence and emergence for a ray of light travelling through a prism in a plane normal to its refracting edge prove that

$$\sin^2 \tfrac{1}{2}(\alpha + \psi) = \sin^2 \tfrac{1}{2}\alpha + \frac{(\mu^2 - 1) \sin^2 \frac{1}{2}\alpha}{1 - \sin^2 \frac{1}{2}(\phi_1 - \phi_2) \sec^2 \frac{1}{2}\alpha},$$

where α, ψ and μ have their usual significance.

Let θ_1 and θ_2 be the angles made by the ray within the prism with

the normals at the points of incidence and emergence of the transmitted ray. If ψ is the angular deviation, we have, in the usual way

$$\left.\begin{array}{l}\sin\phi_1 = \mu\sin\theta_1\\ \sin\phi_2 = \mu\sin\theta_2\end{array}\right\} \quad . \quad (i) \qquad\qquad \left.\begin{array}{l}\phi_1 + \phi_2 = \alpha + \psi\\ \theta_1 + \theta_2 = \alpha\end{array}\right\} \quad . \quad (ii)$$

Adding equations (i),

$$\sin\phi_1 + \sin\phi_2 = \mu[\sin\theta_1 + \sin\theta_2],$$

i.e. $\sin\tfrac{1}{2}(\phi_1 + \phi_2)\cos\tfrac{1}{2}(\phi_1 - \phi_2) = \mu\sin\tfrac{1}{2}(\theta_1 + \theta_2)\cos\tfrac{1}{2}(\theta_1 - \theta_2)$.

Substituting from (ii) and squaring

$$\sin^2\tfrac{1}{2}(\alpha + \psi)\cos^2\tfrac{1}{2}(\phi_1 - \phi_2) = \mu^2\sin^2\tfrac{1}{2}\alpha\cos^2\tfrac{1}{2}(\theta_1 - \theta_2) \quad . \quad (iii)$$

Multiplying equations (i) together, we have

$$\sin\phi_1\sin\phi_2 = \mu^2\sin\theta_1\sin\theta_2,$$

i.e. $\cos(\phi_1 + \phi_2) - \cos(\phi_1 - \phi_2) - \mu^2[\cos(\theta_1 + \theta_2) - \cos(\theta_1 - \theta_2)]$.

$$\therefore \cos^2\tfrac{1}{2}(\phi_1 + \phi_2) - \cos^2\tfrac{1}{2}(\phi_1 - \phi_2)$$
$$= \mu^2[\cos^2\tfrac{1}{2}(\theta_1 + \theta_2) - \cos^2\tfrac{1}{2}(\theta_1 - \theta_2)],$$

i.e. $\mu^2\cos^2\tfrac{1}{2}(\theta_1 - \theta_2) = \mu^2\cos^2\tfrac{1}{2}\alpha - \cos^2\tfrac{1}{2}(\alpha + \psi) + \cos^2\tfrac{1}{2}(\phi_1 - \phi_2)$,

and hence from (iii)

$$\sin^2\tfrac{1}{2}(\alpha + \psi)\cos^2\tfrac{1}{2}(\phi_1 - \phi_2)$$
$$= \sin^2\tfrac{1}{2}\alpha[\mu^2\cos^2\tfrac{1}{2}\alpha - \cos^2\tfrac{1}{2}(\alpha + \psi) + \cos^2\tfrac{1}{2}(\phi_1 - \phi_2)].$$

$$\therefore \sin^2\tfrac{1}{2}(\alpha + \psi)[\cos^2\tfrac{1}{2}(\phi_1 - \phi_2) - \sin^2\tfrac{1}{2}\alpha]$$
$$= \sin^2\tfrac{1}{2}\alpha[\mu^2\cos^2\tfrac{1}{2}\alpha - \sin^2\tfrac{1}{2}(\phi_1 - \phi_2)].$$

$$\therefore \sin^2\tfrac{1}{2}(\alpha + \psi)[-\sin^2\tfrac{1}{2}(\phi_1 - \phi_2) + \cos^2\tfrac{1}{2}\alpha]$$
$$= \sin^2\tfrac{1}{2}\alpha[\cos^2\tfrac{1}{2}\alpha - \sin^2\tfrac{1}{2}(\phi_1 - \phi_2) + (\mu^2 - 1)\cos^2\tfrac{1}{2}\alpha].$$

The required relation now follows.

EXAMPLES XX

1.—State the laws of reflexion and refraction. How would you proceed to verify them experimentally?

2.—A glass block is 5 cm. thick. It is silvered on the back surface, which is parallel to the front surface. A ray of light is incident at an angle of 46°. Trace the ray of light which first emerges from the block whose refractive index is 1·52. What is the angle at which the ray impinges upon the silvered surface?

3.—A ray of light is incident at an angle of 57° upon a plate of glass. The angle of refraction is 31·5°. Find by drawing, and by calculation, the refractive index of the material.

4.—What do you understand by the term critical angle? Calculate the critical angle for water whose μ is 1·334. Water is placed upon a block of glass (μ = 1·500). What is the critical angle for light passing from the glass to the water?

5.—A prism has an angle of 61° 30′. If the angle of minimum deviation is 48° 45′, calculate the refractive index of the material of the prism. What is the critical angle for light travelling from such a material to air?

6.—A block of glass is 5·872 cm. thick. A small speck of dirt on its lower surface is viewed from above. The spot appears to be 2·031 cm. nearer. Calculate the refractive index for this glass and the angle of minimum deviation for a prism made of similar glass, if the angle of the prism is 60°.

7.—Explain why several images of a candle flame may be seen by a person holding a lighted candle in front of a thick mirror. Discuss the difference of the intensity of the images thus formed as the angle of incidence of the light increases.

8.—State the *laws of refraction of light* and explain what is meant by the term *critical angle*. An inch cube is constructed of a material whose index of refraction is 1·65. Calculate the least radius of the opaque circular discs which must be placed centrally over each face of the cube, so that a small air bubble at its centre shall be invisible from an external point.

9.—A small object is placed 20 cm. in front of a block of glass, the remote side of the block being silvered. Determine the position of the image when viewed along a normal to the front surface of the block and passing through the object itself. Thickness of glass = 10 cm. and its index of refraction = 1·52.

10.—Describe and give the theory of an accurate method of determining the refractive index of water. The refractive index for water is $\frac{4}{3}$; for glass it is $\frac{3}{2}$. A ray of light travelling in water is incident at an angle of 40° upon a plane water-glass interface. Calculate the angle of refraction.

11.—Define the terms *refractive index* and *critical angle*, and deduce the relation between the two for any given medium. A metal tank is completely filled with liquid, having a mean refractive index 1·6. A thin circular cork mat is to be floated centrally over a luminous point 6 cm. below the level of the liquid. Calculate the least radius of a mat sufficient to prevent the luminous point from being observed from a point outside the tank.

12.—A glass whose refractive index is 1·652 for sodium light is to be used to construct a prism such that the angle of minimum deviation for such light shall be equal to the angle of the prism. What is the angle of the prism ?

13.—Show that the ray of light which enters the first face of a prism at grazing incidence is least likely to suffer total internal reflexion at the second face. Find the least value of the refracting angle of a prism made of glass of refractive index $\frac{7}{4}$ such that no rays incident on one of the faces containing this angle can emerge from the other face.—(N.H.S.C. '29.)

14.—If ϕ_1 and ϕ_2 are the angles of incidence and emergence for a ray of light travelling through a prism in a plane at right angles to the edge of the prism, show that

$$\sin^2\tfrac{1}{2}(a + \psi) = \sin^2\tfrac{1}{2}a + \frac{(\mu^2 - 1)\sin^2\tfrac{1}{2}a}{1 - \sin^2\tfrac{1}{2}(\phi_1 - \phi_2)\sec^2\tfrac{1}{2}a}$$

where μ is the refractive index of the material of the prism, a the angle of the prism, and ψ the deviation of the ray. Use the above equation to show that the deviation is a minimum when $\phi_1 = \phi_2$.

15.—The refractive indices of a material for three rays are μ_1, μ_2 and μ_3 respectively; if the corresponding angles of minimum deviation for a prism of the same material are γ_1, γ_2 and γ_3 respectively, and these are in arithmetical progression, prove that

$$\frac{\sin\tfrac{1}{2}\gamma_2}{\mu_2} = \frac{\sin\tfrac{1}{2}\gamma_1 + \sin\tfrac{1}{2}\gamma_3}{\mu_1 + \mu_3}.$$

CHAPTER XXI

REFRACTION OF LIGHT AT CURVED SURFACES —LENSES

Refraction at a Concave Spherical Surface.—Let APB, Fig. 21·1, be the principal section of a concave surface bounding a medium whose refractive index is μ. Let C be the centre of curvature and O a luminous point on the axis. If OL is a ray of light incident upon AB, it will be refracted along LM, i.e. it becomes bent towards the normal CLN. Let ML produced cut the axis

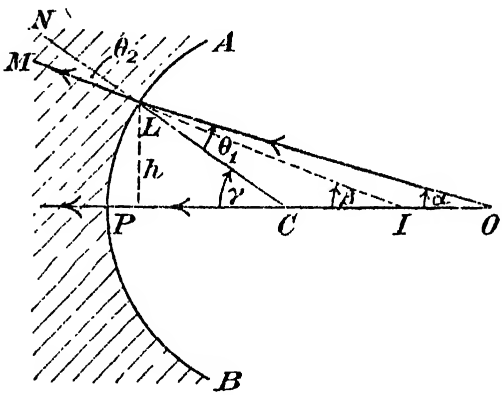


FIG. 21·1.—Refraction of Paraxial Rays at a Convex Spherical Surface.

at I; another ray travels along the axis so that I is the image of O, if it can be shown that all the refracted rays pass through I; let θ_1 be the angle of incidence and θ_2 the angle of refraction which is also equal to the $\widehat{CLI}$; let α, β and γ be the angles shown. If only the small region near P is considered, i.e. the angles of incidence and refraction are small, then, since the sines of small angles are equal to their circular measure, $\theta_1 = \mu\theta_2$, [$\because \sin\theta_1 = \mu \sin\theta_2$].

Now, from the diagram,

$$\theta_1 = \gamma - \alpha, \quad \text{and} \quad \theta_2 = \gamma - \beta.$$

$$\therefore \gamma - \alpha = \mu(\gamma - \beta).$$

Again since the angles α, β, and γ are small, they may be replaced by their tangents, which are approximately equal to $\frac{h}{\text{PO}}$, $\frac{h}{\text{PI}}$ and $\frac{h}{\text{PC}}$ or $\frac{h}{u}$, $\frac{h}{v}$, $\frac{h}{r}$ respectively, where h is the perpendicular distance of L from PO, u and v are the distances of the object and image from P, and r is the radius of curvature of the surface.
Then

$$\frac{1}{r} - \frac{1}{u} = \frac{\mu}{r} - \frac{\mu}{v},$$

or

$$\frac{\mu}{v} - \frac{1}{u} = \frac{\mu - 1}{r}.$$

Refraction at a Convex Spherical Surface.—The path of a ray refracted at a convex surface is indicated in Fig. 21·2. With the same notation as before, we have,

$$\theta_1 = \alpha + \gamma, \text{ and } \theta_2 = \gamma - \beta.$$

Hence

$$\alpha + \gamma = \mu(\gamma - \beta).$$

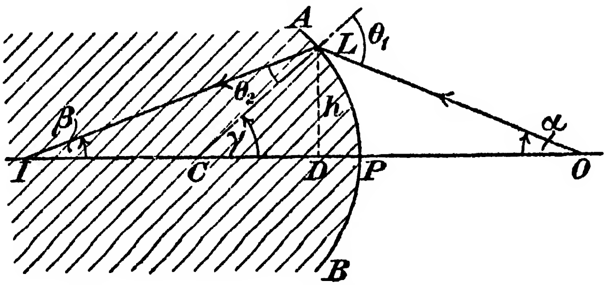


FIG. 21·2.—Refraction of Paraxial Rays at a Convex Spherical Surface.

If α, β, and γ are small, we may replace them by their respective tangents, and remembering that v and r are negative, we have,

$$\frac{1}{u} - \frac{1}{r} = \mu\left(-\frac{1}{r} + \frac{1}{v}\right),$$

i.e.

$$\frac{\mu}{v} - \frac{1}{u} = \frac{\mu - 1}{r}.$$

The General Form of the Equation for Refraction at a Curved Spherical Surface.—Let μ_1 and μ_2 be the refractive indices of the media on the two sides of the curved spherical surface whose radius of curvature is r. Then from Fig. 21·1 we have, $\mu_1 \sin\theta_1 = \mu_2 \sin\theta_2$ for the refraction at L, and this equation assumes the form $\mu_1\theta_1 = \mu_2\theta_2$, since θ_1 and θ_2 are small. But $\theta_1 = \gamma - \alpha$, and $\theta_2 = \gamma - \beta$, so that

$$\mu_1(\gamma - \alpha) = \mu_2(\gamma - \beta).$$

With the same limitations as before, this becomes

$$\mu_1\left(\frac{1}{r} - \frac{1}{u}\right) = \mu_2\left(\frac{1}{r} - \frac{1}{v}\right).$$

This is the equation required, and in the above form is easily remembered. It is known as the ***fundamental paraxial equation*** for a surface.

The same equation is true for refraction at a convex spherical surface—the proof is left as an exercise. The advantage of using this general form of the equation is that difficulties in solving problems based on the following experiment are avoided.

Moreover, if we write $\mu_1 = -\mu_2$ so that the refraction becomes a reflexion [cf. p. 396], we obtain $\frac{1}{r} + \frac{1}{u} = \frac{2}{r}$, which is the well-known relation concerning conjugate foci when an image is formed by the reflexion of light rays at a spherical surface.

Experiment.—Fill a thin-walled cylindrical tank with water and place an upright pin in the water. Let the pin be viewed through the curved surface of the tank and the position of the image be located by a parallax method. The radius of the tank is measured. The refractive index of the water is calculated as follows.

Distance of pin from pole of surface	= 18·3 cm.
Distance of image from pole of surface	= 21·5 cm.
Radius of curvature of surface	= 12·3 cm.

Then μ_1 is the refractive index of water, while $\mu_2 = 1$, since the medium into which the refraction occurs is air. Hence,

$$\mu_1\left[\frac{1}{12\cdot 3} - \frac{1}{18\cdot 3}\right] = 1\left[\frac{1}{12\cdot 3} - \frac{1}{21\cdot 5}\right],$$

giving $\mu_1 = 1\cdot 31$.

Refraction through a Lens.—A lens is defined as a portion of a transparent refracting medium bounded by two surfaces which are generally spherical or cylindrical. Lenses are divided into two classes; those which are thicker at the centre than at the periphery are termed *convex* or *converging*; those which are thinner are *concave* or *diverging*. The more simple types of lenses are indicated in Fig. 21·3.

Suppose that μ is the index of refraction of the medium of a lens with respect to air, and that u is the distance of a luminous object from the nearer surface of the lens whose radius is r_1—the object is assumed to be on the optical axis of the lens. If v' is the distance from this surface at which the image would be formed, if the second face were absent, then

$$\frac{\mu}{v'} - \frac{1}{u} = \frac{\mu - 1}{r_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

But this image will serve as an object when refraction takes place at the second face, i.e. we now have an object at distance $(v' + t)$

from this second face, if t is the thickness of the lens. Let r_2 be the radius of curvature of this second face, then, if the final image is at distance v from this face.

$$\frac{\frac{1}{\mu}}{v} - \frac{1}{(v'+t)} = \frac{\frac{1}{\mu}-1}{r_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The value $\frac{1}{\mu}$ is used for the refractive index, since the refraction

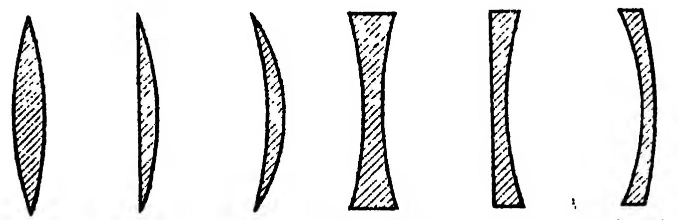


FIG. 21·3.—Different Types of Thin Lenses.

takes place from glass to air. If t is small, so that it can be neglected, then the above equation may be written

$$\frac{1}{v} - \frac{\mu}{v'} = \frac{1-\mu}{r_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Adding (1) and (3) in order to eliminate v', we have

$$\frac{1}{v} - \frac{1}{u} = (\mu - 1)\left[\frac{1}{r_1} - \frac{1}{r_2}\right].$$

When the object is at such a point on the axis that the image is at infinity, i.e. $v = \infty$, the object is said to be at the ***first principal focus*** of the lens, while the distance of the object from the lens is known as its ***first focal length***, f_1. This is given by

$$-\frac{1}{f_1} = (\mu - 1)\left[\frac{1}{r_1} - \frac{1}{r_2}\right]$$

If the object is at an infinite distance from the lens, $\frac{1}{u} = 0$ and the image is formed at a distance f_2 given by

$$\frac{1}{f_2} = (\mu - 1)\left[\frac{1}{r_1} - \frac{1}{r_2}\right]$$

This particular distance denoted by f_2 is termed the ***second focal length*** of the lens, the point on the axis at which the image is formed being the ***second principal focus*** of the lens.

From these equations we see that for thin lenses the two focal lengths of a lens are numerically equal but that the two principal foci are on opposite sides of the lens. ***In the sequel when we speak of the focal length of a lens we shall always imply its second focal length.***

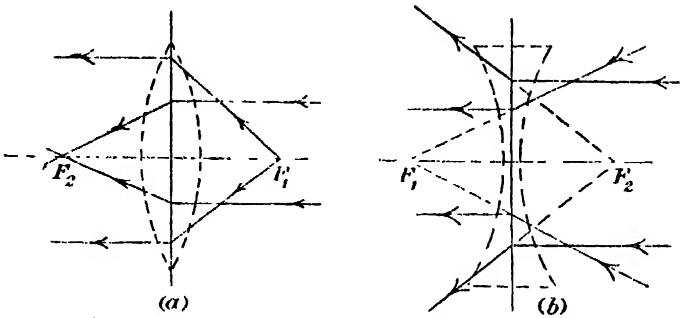


FIG. 21·4.—The Focal Points of (a) Converging and (b) Diverging Lenses. [Cf. footnote on p. 383.]

The paths of rays of light proceeding from the first principal focus of a converging lens and to its second principal focus are shown in Fig. 21·4 (a). The case of a diverging lens is treated in Fig. 21·4 (b). Here, it should be noted that the rays do not actually pass through the focal points.

The Action of Lenses.—To explain the action of lenses let us refer to Fig. 21·5 (a) and (b). In (a) we have two series of truncated prisms of different angles arranged symmetrically with reference to an axis and with their bases parallel to this axis. Consider a luminous point source at O on the above axis. Then a ray of light such as OA is deviated by the prism on which it falls. Now the greatest deviation will be produced by the prism

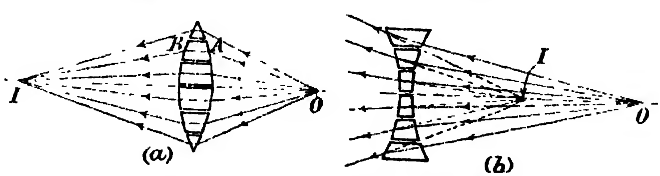


FIG. 21·5.—The Action of Lenses.

farthest from the axis. Such a system of prisms tends to make all the rays converge. If the number of prisms is increased indefinitely, their heights suffering a corresponding diminution, the system approximates to a double convex lens. In the same way a double concave lens may be regarded as an infinite array of

such prisms having their bases turned away from the axis—cf. Fig. 21·5 (*b*). Any diverging beam of light falling on such a system is made more divergent and the emergent rays appear to come from a point on the same side of the system as is the object.

Optical Centre of a Lens.—Let C, C_1, Fig. 21·6, be the centres of curvature of the two faces of a double convex lens, so that CC_1 is the principal axis. Through C draw any radius CR, and through C_1 draw a parallel radius C_1Q. Let PQRS be the path of a ray through the lens. The ray PQ is parallel to RS since the normals at Q and R are parallel to each other. It is required to calculate the position of O, the point in CC_1 at which QR crosses it. The Δ's ORC and C_1QO are similar.

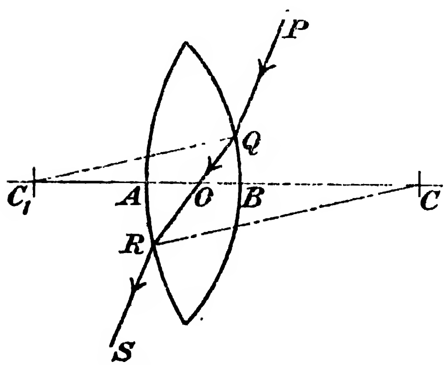


FIG. 21·6.—The Optical Centre of a thin Lens.

$$\therefore \frac{OC}{OC_1} = \frac{CR}{C_1Q} = \frac{CA}{C_1B} = \frac{CA - OC}{C_1B - OC_1} = \frac{OA}{OB}.$$

Hence the position of O is invariable, i.e. it is independent of the choice of R, and it is called the ***optical centre of the lens***; it is characterized by the fact that all rays which pass through it leave the lens parallel to their original direction. For the ***thin*** lenses which are here discussed this optical centre is the same as the mid point of the lens; rays passing through this point are not deviated.

Graphical Construction of the Images of Finite Objects formed by Lenses.—Let OA, Fig. 21·7 (*a*), be a small finite object lying in a plane normal to the principal axis of a lens and being at a distance from the lens greater than its focal length. A ray AD parallel to the axis passes after refraction through F_2 the second principal focus of the lens. The ray AC through the centre of the lens is not deviated so that the intersection of these two rays gives the position of the image of A. Since the ray OC passes along the axis of the lens the image of OA is obtained by drawing BI perpendicular to the axis.

A similar construction has been made in Fig. 21·7 (*b*), where the object is nearer to the lens than is F_1. In this instance the refracted

rays DF_2 and AC never actually intersect but only appear to come from B, a point on the same side of the lens as is the object. The image is a virtual one.

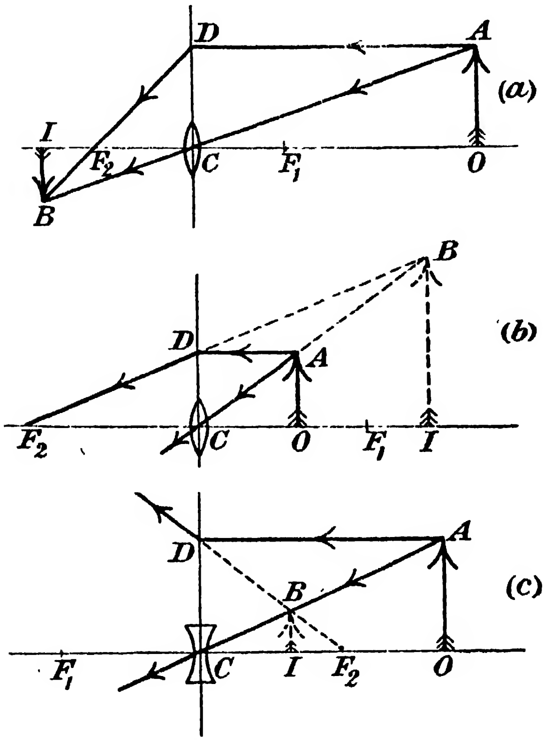


FIG. 21·7.—Graphical Construction of Images formed by Lenses.

The appropriate construction for the image formed by refraction through a concave lens is indicated in Fig. 21·7 (c). Here it must be noted that F_2 is on the same side of the lens as is the object and that the image is virtual.

The Tracing of Pencils of Rays through a Lens.—The position and size of the image having been determined, the course of the rays by which an eye placed near to the axis sees the image may be shown as follows :—In Fig. 21·8 the positions and sizes of the object and image formed in the first instance discussed above have been redrawn. If E is an eye, by joining the extremities of the pupil to B and producing these lines to cut the principal plane of the lens in H and K we obtain the confines of the refracted pencil of light by which the eye observes the point B in the image. If H and K are

joined to A we have the complete pencil from A to E. Similarly the pencil from O to E is constructed.

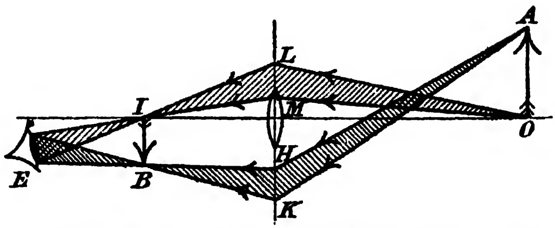


FIG. 21·8.—Method of Tracing Rays through a Lens.

Conjugate Foci.—An inspection of Fig. 21·7 (*a*) shows that if OA is the object then IB is the image, whereas if IB is the object then OA is the image. The points O and I are termed ***conjugate foci.***

Referring to Figs. 21·7 (*b*) and (*c*) the points O and I are conjugate foci in the sense that if rays of light forming an image at IB in the absence of the lens are incident upon a lens at C, then a real image will be produced at OA.

Focal Planes and Secondary Axes.—Planes drawn at right angles to the principal axis of a lens and passing through its principal focal points are termed the ***first and second focal planes*** of the lens. A straight line through C, the lens centre, is called a ***secondary axis*** of the lens.

More About the Tracing of Rays through a Lens.—Suppose that OA, Fig. 21·9 (*a*), is a ray of light incident upon a converging lens whose principal foci are F_1 and F_2. Then planes through these points perpendicular to the principal axis of the lens are the focal planes. The path of the ray OA after refraction through the lens is required. Let OA cut the first focal plane in H. If C is the centre of the lens, a ray HC would pass undeviated through the lens. Now a diverging pencil of light AHC, originating from a point H in the first focal plane of the lens will emerge as a parallel beam of light after refraction. Since HC produced, i.e. HCE, is one ray of this parallel beam the ray OA must be refracted along AB, where AB is parallel to HCE. [HCE is known as a ***secondary axis*** of the lens.]

In the above we have made use of a special property of the first focal plane of a converging lens. We could have used the second focal plane equally well. Thus if JC, Fig. 21·9 (*b*), is a ray parallel to OA which, when produced, cuts the second focal plane in K, then a parallel beam of light OACJ falling on the lens will be brought to a focus at K. OA must therefore be refracted along AK.

When the lens is a diverging one, let OA, Fig. 21·9 (*c*), be an incident ray which, when produced, cuts the first focal plane in H. Now a ray DC incident at C and travelling in the direction DCH passes through the lens without deviation. A cone of light OHD,

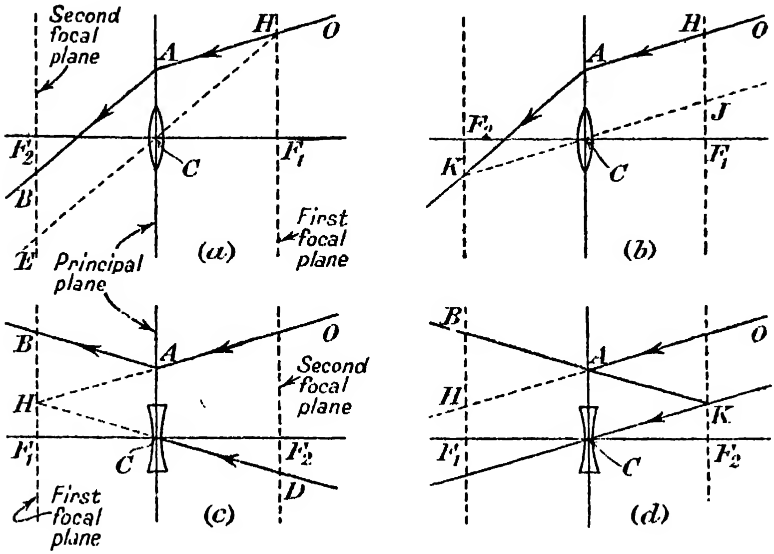


FIG. 21·9.—The Chief Properties of Focal Points and of Focal Planes.

however, converging on H, emerges as a parallel beam since H lies in the first focal plane of the lens. Since CH is one ray in this beam the refracted ray corresponding to OA is AB, where AB is parallel to CH.

Fig. 21·9 (*d*) shows how to proceed if the second focal plane of the lens is used.

The Focal Length of a Lens Combination.—When two thin lenses are placed in contact they can be regarded as a single lens. Let f be the focal length of the combination, i.e. the focal length of a single lens having optical properties equivalent to those of the two lenses in contact, whilst f_1 and f_2 are the focal lengths of the constituent lenses. If u is the distance of an object from the combination of lenses, an image will be produced *by the first component* at a distance v' where

$$\frac{1}{v'}-\frac{1}{u}=\frac{1}{f_1}.$$

This image can be regarded as an object with respect to the second

lens which gives rise to an image at distance v from the system. Then

$$\frac{1}{v}-\frac{1}{v'}=\frac{1}{f_2},$$

or, by addition,

$$\frac{1}{v}-\frac{1}{u}=\frac{1}{f_1}+\frac{1}{f_2}.$$

Now if the combination is replaced by a single lens which gives an image of an object at distance u at a distance v, then the focal length, f, of this lens is given by

$$\frac{1}{f}=\frac{1}{v}-\frac{1}{u}.$$

Hence

$$\frac{1}{f}=\frac{1}{f_1}+\frac{1}{f_2},$$

i.e. the power of the combination is the sum of the powers of the constituent lenses. [The ***power*** of a lens is defined as the reciprocal of its focal length expressed in metres. The unit of power is the ***dioptre***.]

Linear Magnification.—The linear magnification of a lens is defined as the ratio of the size of the image to that of the object and will always be considered positive. Referring to Fig. 21·7 (*a*) we see that

$$m=\left|\frac{\mathrm{IB}}{\mathrm{OA}}\right|=\left|\frac{v}{u}\right|$$

since the triangles AOC and BIC are similar.

Minimum Distance between Image and Object [Converging Lens].—In attempting to arrange a convex lens to produce a real image on a screen, much time is often lost because it is not realized that unless the object and screen are at a distance apart greater than a certain minimum value it is impossible to obtain an image on the screen. To calculate this minimum distance in terms of the focal length of the lens, let U, V, and F denote the numerical values of the quantities u, v, and f respectively. Then

$$\frac{1}{\mathrm{V}}+\frac{1}{\mathrm{U}}=\frac{1}{\mathrm{F}},$$

and we have to determine the minimum value of (U + V) subject to the above condition, i.e. to

$$\mathrm{UV}=(\mathrm{U}+\mathrm{V})\mathrm{F}.$$

Hence $$\mathrm{U^2V^2}=(\mathrm{U^2}+2\mathrm{UV}+\mathrm{V^2})\mathrm{F^2},$$

and $$UV(UV - 4F^2) = (U^2 - 2UV + V^2)F^2 = a + ve \text{ quantity}.$$

$$\therefore UV - 4F^2 > 0 \quad [\because UV \text{ is } + ve]$$
$$(U + V)F - 4F^2 > 0$$
$$(U + V) > 4F.$$

Hence the minimum distance is 4F.

Alternative Proof. Let $Y = U + V$: then the minimum value of Y, subject to the condition that $\frac{1}{V} + \frac{1}{U} = \frac{1}{F}$, has to be determined. Eliminating V from these equations, we have

$$Y = U + \frac{UF}{U - F}.$$

Hence $$\frac{dY}{dU} = 1 + \frac{F(U - F) - UF}{(U - F)^2} = \frac{U(U - 2F)}{(U - F)^2}.$$

For a minimum (or maximum) $\frac{dY}{dU} = 0$, so that $U = 0$ or $U = 2F$.

The solution $U = 0$ is of no physical importance for it implies that the object is in contact with the lens.

The solution $U = 2F$ means that when this is fulfilled the distance between object and image is a minimum or a maximum.

A second differentiation gives $\frac{d^2Y}{dU^2} = \frac{2F^2}{(U - F)^3}$, and when $U = 2F$, this is $\frac{2}{F}$, a positive quantity. Thus Y is a minimum, and since $U = V$ when $U = 2F$, the minimum distance required is 4F.

The Equation $pq = -f^2$.—Suppose that P, Fig. 21·10 (*a*), is the pole of a lens whose principal foci are F_1 and F_2 respectively.

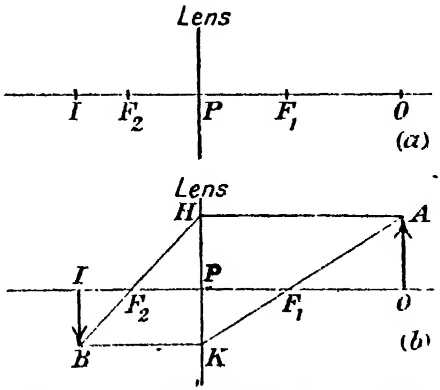


Fig. 21·10.—The Equation $pq = -f^2$.

Let O and I be two conjugate points on the principal axis of the above lens. Then, with the usual algebraic convention with respect to signs,

$$u = PO, \text{ and } v = -IP = PI.$$

Moreover, if, as usual, f is the second focal length of the lens, then

$$f = -F_2P = PF_2 = -PF_1.$$

The formula $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$ may therefore be written

$$-\frac{1}{IP} - \frac{1}{PO} = -\frac{1}{F_2P},$$

i.e.
$$-\frac{1}{IF_2 + F_2P} - \frac{1}{PF_1 + F_1O} = -\frac{1}{F_2P}.$$

Let $p = F_1O$ and $q = F_2I$. Then

$$-\frac{1}{(-q-f)} - \frac{1}{(-f+p)} = +\frac{1}{f}.$$

Whence
$$pq = -f^2.$$

This equation is sometimes known as *Newton's equation,* since Newton first obtained it.

Students who have difficulty with the above analytical proof may find the following geometrical proof for a converging lens instructive. The diagram shown in Fig. 21·10 (*b*) refers to the formation of a real image by the lens. All distances will be considered numerically. Then from the similar triangles IBF_2 and F_2HP, we have

$$\frac{IF_2}{IB} = \frac{F_2P}{HP}.$$

Similarly,
$$\frac{F_1O}{OA} = \frac{PF_1}{PK}.$$

Hence
$$F_1O.IF_2 = F_2P.PF_1.$$

i.e.
$$|pq| = |f|^2.$$

Worked Examples.—(i) Calculate the refractive index of the material of a converging lens of focal length 15 cm., the radii of curvature of its faces being 20 cm. and 12 cm. respectively.

The focal length is negative. Suppose that the 20 cm. face is nearer to the object. Then this is r_1 and is negative; r_2 is positive.

Since
$$\frac{1}{f} = (\mu - 1)\left[\frac{1}{r_1} - \frac{1}{r_2}\right],$$

we have
$$-\frac{1}{15} = (\mu - 1)\left[-\frac{1}{20} - \frac{1}{12}\right]$$

$$\therefore \mu = 1{\cdot}50.$$

(ii) What lens must be placed in contact with a diverging lens of focal length 25 cm. in order that the lens combination may produce a real image magnified 3 times of an object 20 cm. from the combination.

Let v be the distance of the image from the lens combination.

Since the image is real, v is negative, and $v = -3u$, since $m = \frac{|v|}{|u|}$

Hence $$v = -60 \text{ cm.}$$

$$\therefore \frac{1}{\phi} = -\frac{1}{60} - \frac{1}{20} = -\frac{1}{15}.$$

But $\frac{1}{\phi} = \frac{1}{f_1} + \frac{1}{f_2}$ where f_1 is $+25$ cm. and f_2 is to be found.

$$\therefore -\frac{1}{15} = +\frac{1}{25} + \frac{1}{f_2}, \qquad \therefore f_2 = -9{\cdot}4 \text{ cm.}$$

A converging lens of focal length 9·4 cm. is required.

(iii) A luminous point on the axis of a symmetrical biconvex lens of focal length 100 cm. appears to be at the centre of curvature of the second face of the lens when viewed through the lens. If the object is 55 cm. from the lens, calculate the refractive index of its material.

$$f = -100 \text{ cm.} \qquad u = +55 \text{ cm.} \qquad \text{What is } v\,?$$

$$\frac{1}{v} - \frac{1}{55} = -\frac{1}{100},$$

$$\therefore v = \frac{5{,}500}{45} \text{ cm.} \qquad \text{This is } r_2. \qquad \text{Hence } r_1 = -\frac{5{,}500}{45} \text{ cm.}$$

$$-\frac{1}{100} = (\mu - 1)\left[-\frac{45}{5{,}500} - \frac{45}{5{,}500}\right]$$

$$\therefore \mu = 1{\cdot}61.$$

EXAMPLES XXI

1.—For a converging lens $u = 81{\cdot}6$ cm., $v = -30{\cdot}4$ cm. What is f ? Draw a figure to check your calculation.

2.—For a diverging lens $u = 20{\cdot}6$ in., and $f = 14{\cdot}2$ in. Where is the image, and what is its magnification ?

3.—The glass of a thin converging lens has a refractive index 1·51. Its focal length in air is 10·3 cm. What is its focal length when placed in water whose refractive index is 1·34 ?

4.—Two converging lenses each have a focal length 15 cm. They are 30 cm. apart. An object is placed 10 cm. in front of the first lens. Where is the image seen through the second lens ? What is its magnification ?

5.—A candle is placed 76·3 cm. from a screen. There are two positions in which a converging lens can be placed so that an image of the candle appears on the screen. These positions are 20·5 cm. apart. What is the focal length of the lens ?

6.—A converging lens of focal length 12·7 cm. is placed in contact with a concave lens. The whole is equivalent to a converging lens of focal length 18·4 cm. What is the focal length of the diverging lens ?

7.—The focal length of a converging lens is 8·2 cm. This is placed 12·1 cm. in front of a concave mirror whose radius of curvature is 5·4 cm. Determine the size and position of an image of an object 3 cm. high placed 3·7 cm. in front of the lens, the image being produced

by refraction through the lens, and then by reflexion at the concave mirror. Check by a diagram.

8.—When a pin is placed at a distance a from a converging lens an image is obtained at a distance b from the lens. Draw a rough graph to indicate the form of the relation between a and $a + b$ for different values of a, for virtual as well as for real images. Show how the focal length of the lens may be deduced from the graph.

9.—How would you combine a converging lens and a plane mirror so as to give an image of a pin coincident with the object and (*a*) erect, (*b*) inverted? In each case, give a diagram showing how the image is formed, and explain how the experiment enables the focal length of the lens to be determined.

10.—A converging lens floats on mercury. A pin and its image appear to coincide when the pin is 10·3 cm. from the lens. If the lens has a focal length 20·6 cm. what is the radius of curvature of the lens surface in contact with the mercury?

11.—A converging meniscus lens having a focal length of 22·5 cm. is held in front of illuminated cross-wires. Images appear in turn at the side of the cross-wires when the lens is 7·8 cm. and 3·9 cm. away from the wires. Calculate the refractive index of the material of the lens.

12.—A glass sphere is 10 cm. in diameter. A small air bubble inside the sphere appears to be 2 cm. from the nearer surface of the sphere when it is viewed along that line which passes through the bubble and the centre of the sphere. What is the true position of the bubble if the refractive index of the material of the sphere is 1·5?

13.—A converging meniscus lens is set up with a plane mirror behind it, the axis of the lens being normal to the mirror. A white screen with a small illuminated aperture in it is placed some distance in front of the lens and gradually moved up towards it. It is found that images of the aperture are focused on the screen when it is at the following distances from the lens:—35 cm., 25 cm., and 8 cm. The image at 35 cm. disappears when the mirror is removed. Calculate the focal length, the radii of curvature of the surfaces, and the refractive index of the material of the lens.

CHAPTER XXII

THE PRACTICAL DETERMINATION OF THE OPTICAL CONSTANTS OF MIRRORS AND LENSES

The Location of Images.—The position of the image of a pin formed by an optical instrument may be found by a parallax method. To explain this method let E, Fig. 22·1, be the eye of an observer when viewing two pins, P_1 and P_2, along the straight line joining them. The two images will be superimposed on the retina and, in general, it will be impossible to decide which is the nearer pin. To ascertain this fact the eye is moved slightly to one side into a position E_1 or E_2 when the images on the retina will no longer coincide. The more distant object P_2 will apparently move to the same side of P_1 as does the observer. We say that ***parallax*** exists between the two pins.

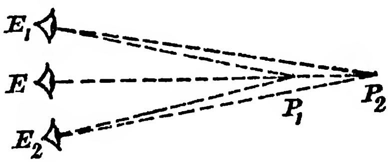


FIG. 22·1.

The same argument applies if P_1 is a pin and P_2 the image of another pin. Hence, if these are in such a position that there is no parallax between them it follows that P_1 and P_2 must coincide.

The Optical Bench.—A convenient piece of apparatus for use in many optical experiments is the optical bench, which consists essentially of a long, straight, rigid bar of metal graduated in cm., etc. A number of stands to hold various pieces of apparatus may be moved along the bar. One stand carries a piece of cardboard across a hole in which there is placed a piece of wire gauze. When this is illuminated by a lamp it serves as an object. Other stands carry the lens or mirror and a screen to receive the real image. Before attempting any work with an optical bench the various pieces of apparatus must be adjusted so that their centres are coaxial. The distance between the centres of the apparatus carried in any two stands is determined with the aid of a measuring rod fixed horizontally in one of the stands. If possible, the same end of this rod is brought in turn into contact with the centres of the two objects whose distance apart is required. The difference of the

two readings indicated by the pointer attached to the stand carrying the rod gives the required distance. Often, it is more expedient to bring opposite ends of the rod into contact with the two objects. When this is so the distance required is the sum of the displacement of the stand carrying the rod and its length.

The Radii of Curvature and Focal Lengths of Concave Surfaces.—***Method i:*** We have already seen that the image of an object in a plane through the centre of curvature of a concave surface and normal to its axis lies in that plane and that it is equal in size to the object but inverted. If, therefore, a concave mirror is placed in front of an illuminated piece of wire gauze and moved until a sharp image is formed immediately below the gauze, the distance between the mirror and the gauze is equal to the radius of curvature of the mirror.

Method ii: The illuminated wire gauze is placed slightly above the axis of the mirror and the image, which is then formed just below the axis, obtained on a white screen. The distances u and v may be measured and f calculated. A series of observations should be taken and a mean value of f deduced.

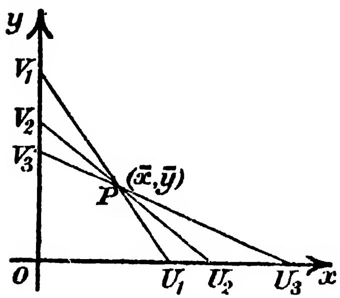


Fig. 22.2.

A mean value may also be deduced graphically as follows:—Points U and V on rectangular axes Ox and Oy and such that OU $= u$, OV $= v$, are plotted (due attention being paid to signs) and a straight line drawn through them. The equation to this line is

$$\frac{x}{u}+\frac{y}{v}=1.$$

In Fig. 22·2 a series of such lines for corresponding values of u and v are shown. Now the particular values of the intercepts on the two axes made by any one of these lines are related by the equation

$$\frac{f}{v}+\frac{f}{u}=1.$$

The above two equations indicate that the point whose co-ordinates are $(\bar{x}=f, \bar{y}=f)$ lies on this line. Since we have considered the general equation to these lines it follows that they all pass through the point (f, f) shown at P.

The Radii of Curvature and Focal Lengths of Convex Surfaces.—***Method i:*** A long pin AB, Fig. 22·3, is placed in front of a convex mirror M and a plane mirror N, and the eye E of an

observer directed along the axis. Two images I and C will be seen. The one formed by reflexion from the convex mirror consists of a diminished and virtual image of the upper part of the pin, while that formed by the plane mirror is a virtual image of the lower part of the pin. Its magnification is unity, and it is at the same distance behind the mirror N as the object is in front of it. The

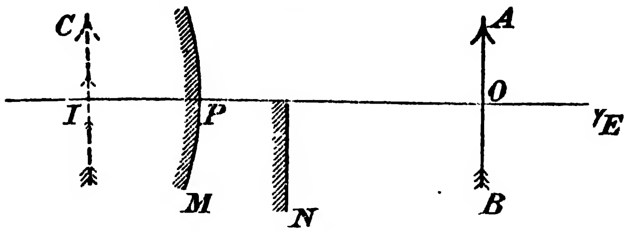


FIG. 22·3.

mirror N is moved until there is no parallax between the two images. The appropriate values of u and v may be determined from the positions of the various pieces of apparatus. Instead of working out each of a series of observations, a graphical method similar to the above may be used. It will be found that $(\bar{x}, \bar{y})$ lies in the third quadrant, i.e. each co-ordinate is negative.

Method ii: In this method an auxiliary convex lens is used to form a real image of a small object. The convex surface is then placed behind the lens and moved until an image is produced adjacent to the object. This image is inverted and is formed when the distance between the pole of the convex mirror and the screen

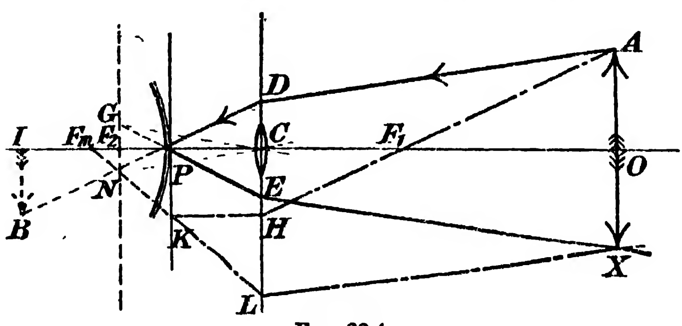


FIG. 22·4.

on which the image IB was obtained is equal to the radius of curvature of the mirror. To show that the image is inverted we shall make an accurate drawing and avoid the difficulties of an analytical proof.

Let OA, Fig. 22·4, be the object and C the centre of a lens whose principal foci are F_1 and F_2. In the absence of the convex mirror

the real image IB is produced. [The construction lines used are not shown.] Let us assume that P is the pole of a convex mirror whose radius of curvature is PI. Consider the ray AD which after refraction through the lens travels in the direction DPB, i.e. towards the pole of the mirror. It is reflected along PE where $D\widehat{P}C = E\widehat{P}C$. To determine the path of this ray after passing through the lens we produce EP to cut the second focal plane of the lens in G and draw the secondary axis GC. Then the refracted ray is EX where EX is parallel to GC. Also consider the ray HK travelling towards the mirror in a direction parallel to the axis of the system. After reflexion it travels along KL, where LK produced passes through F_m, the focus of the mirror. If this line cuts the second focal plane in N, the line LX parallel to the secondary axis NC gives us the refracted ray. Since these two refracted rays intersect at X and the ray OC is reflected back along CO, the image must be OX.

[Attention is again called to the fact that in all such diagrams as Fig. 22·4, the scale in a direction perpendicular to the optical axis is very much enlarged, so that the ray AD does actually pass through the lens, and that the lens in the diagram is only to remind us of the type of lens in use and that it does not represent the lens on the same scale as the rest of the diagram.]

The Focal Lengths of Converging Lenses. ***Method i:*** The lens is arranged to produce a real image of a piece of illuminated gauze and the distances u and v measured. The value of the focal length is then calculated. A graphical method similar to that described for mirrors may also be used. If u and v are the intercepts made by a straight line on the axes OX, OY, its equation is

$$\frac{x}{u} + \frac{y}{v} = 1.$$

Since u, v and f are related by the equation

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$$

it follows that the point $(-f, f)$ lies on all such lines for any one lens. Since f is negative for convex lenses, it follows that the point $(-f, f)$ lies in the fourth quadrant.

Method ii: In this method use is made of the fact that rays of light proceeding from a point in the first focal plane of a convex lens form a parallel beam after refraction through the lens. The direction of this beam is parallel to that secondary axis passing through the luminous point and the centre of the lens. If such a beam falls upon a plane mirror it will be reflected as a beam of parallel light and if this passes through the lens an image will be

produced in the plane containing the object. We now have to show that the image is equal in size to the object but inverted.

Let OA, Fig. 22·5, be a small object (illuminated wire gauze) in the first focal plane of a convex lens whose centre is C. Let AD be a ray which after refraction passes along DE where DE is parallel to the secondary axis AC. This ray will be reflected along EF where DE and EF make equal angles with the normal to the plane mirror at E. If CB the secondary axis parallel to EF is constructed then after refraction the ray EF travels along FB where B is the point of intersection of CB with the focal plane through O. Since the ray OC will be reflected along the principal axis of the lens the image will be OB, where OB is perpendicular to OC. Since the Δ's OAC and OBC are congruent OA = OB.

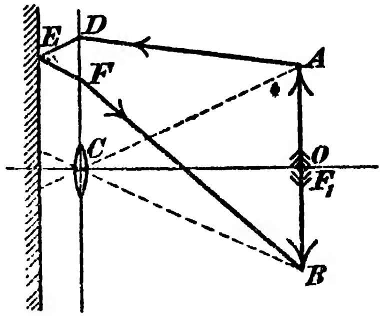


FIG. 22·5.

Method iii: This is known as the ***displacement method.*** The convex lens, Fig. 22·6, is arranged to form a real image of some illuminated cross-wires, P, on a screen S. [The distance between P and S must therefore be greater than $|4f|$—cf. p. 430.] The wires and screen being fixed in position, the lens is moved to a position B such that a real image is again produced on the screen. Now the distance of the lens from the wires at P in the first instance is equal to the distance of the lens from the screen S in the second instance. Hence, using a, b, and c to denote the numerical values of the distances indicated, we have, in the first instance, $|u| = c$ and $|v| = b + c$. Hence,

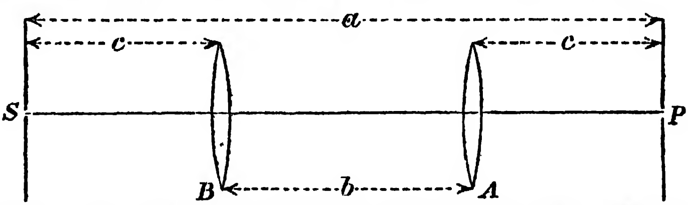


FIG. 22·6.—Focal Length of a Converging Lens by Displacement Method.

$$-\frac{1}{(b+c)} - \frac{1}{c} = -\frac{1}{|f|}.$$

But $a = b + 2c$, so that

$$|f| = \frac{(a^2 - b^2)}{4a}.$$

The great advantage of this method is that we have to measure the shift of the lens and therefore the method does not involve any error due to an incomplete knowledge of the position of the optical centre of the lens under investigation.

Method iv: If the linear magnification and either v or u be known the focal length of the lens may be determined. A slit exactly 1 cm. long is used as object and the image focused on a ground-glass screen having a mm. scale engraved on it. This enables the magnification to be read off at once. The lens, etc., are first arranged so that the magnification is unity. The lens is kept fixed and the slit and scale moved until the magnification is 2, 3, 4, etc. The distance through which the scale is moved is numerically equal to the focal length of the lens. Let v_1 and v_2 be the distances of the image when the magnification is 1 and 2 respectively. Then since $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$ we have $-\frac{1}{|v|} - \frac{1}{|u|} = -\frac{1}{|f|}$. Since the magnification is $\frac{|v|}{|u|}$, we have

$$m = \frac{|v|}{|f|} - 1.$$

When m is respectively 1 and 2, we have,

$$1 = \frac{|v_1|}{|f|} - 1, \text{ and } 2 = \frac{|v_2|}{|f|} - 1.$$

Hence

$$|f| = |v_2| - |v_1|.$$

The expression $|v_2| - |v_1|$ measures the displacement of the screen upon which the image is received, if the lens remains fixed in position.

It is interesting to note that this method may be used to determine the focal length of a thick lens or of a system of lenses.

Alternative procedure: Instead of working with integral values of m, we may determine a series of corresponding values of v and m. Since $m = \frac{|v|}{|f|} - 1$, we may write $y = \frac{1}{|f|}x - 1$, so that $\frac{1}{|f|}$ is the slope of the straight line obtained by plotting $y = m, x = |v|$.

Sometimes the lens whose focal length is required is inaccessible, e.g. it may be 'hidden' in a tube with thin glass covers over the ends. In such a case the distance of the image may be measured from any convenient point, e.g. from the centre of the nearer cover glass. Let it be $|V|$. Then $|v| = |V| + |a|$, where $|a|$ is an unknown constant. The above equation then becomes

$$m = \frac{|V| + |a|}{|f|} - 1$$

$$= \frac{|V|}{|f|} + \text{constant}.$$

If $m = y$ and $|V| = x$, the above equation becomes

$$y = \frac{1}{|f|}x + \text{const.}$$

This represents a straight line whose slope is $\frac{1}{|f|}$. A series of observations on m and corresponding values of $|V|$, enable $|f|$ to be found.

Method v: The following method is due to the late Professor SILVANUS P. THOMPSON. The apparatus required is the same as in the last experiment, but the scale is fixed one metre from the lens and the slit (1 cm. long) moved until a clear image is formed. Let the image be m cm. long. Then the linear magnification is m, and we have

$$m = \frac{100}{|f|} - 1, \text{ or } |f| = \frac{100}{1+m}.$$

Method vi: A telescope is adjusted so that a clear image of some distant object is formed. The convex lens whose focal length is required is then placed coaxially in front of the telescope and a piece of printed matter held at a little distance from the lens. The plane of the paper must be normal to the common axis of the lens and telescope. The distance of the paper is altered until a clear image of the print is seen through the telescope. Since the telescope has been adjusted for parallel light it follows that the printed paper must be in the first focal plane of the lens. The distance between the lens and paper is therefore $|f|$.

The Focal Lengths of Diverging Lenses.—*Method i:* The lens is placed in contact with a converging lens of known focal length and of such power that the combination is equivalent to a convex lens. The focal length of the combination is determined. Let this be f. Then if f_1 and f_2 are the [second] focal lengths of the convex and concave lenses respectively, $\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$, so that f_2 may be calculated.

Method ii: If the available convex lens C, when placed in contact with the concave lens D, is not sufficiently powerful for the combination to be convergent, the lens C is arranged to form a real image I_1 of an object O, Fig. 22·7. The concave lens is then placed between the convex lens and I_1 and a real image I_2 produced. Then I_1 is acting as a virtual object giving a real image I_2. If u and v are respectively the distances of the 'object' I_1 and the image I_2 from the centre of the concave lens, its focal length may be calculated. If a series of corresponding values of u and v are obtained a graphical method similar to that used on p. 438 may be employed.

It is left as an exercise to the student to show that the lines intersect at the point $(-f, f)$ lying in the second quadrant, since f is positive.

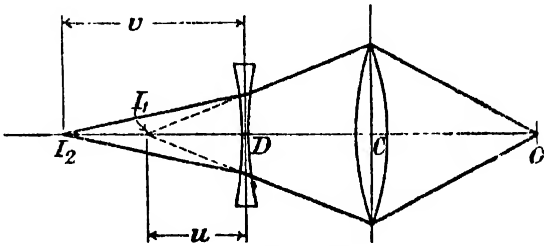


FIG. 22·7.

Method iii. A convex lens L_1, Fig. 22·8, is arranged to produce a real image I of a point source O situated on the axis CO of the lens. The concave lens L_2 and a plane mirror are then placed behind the lens L_1 and moved to such a position that an image of the object

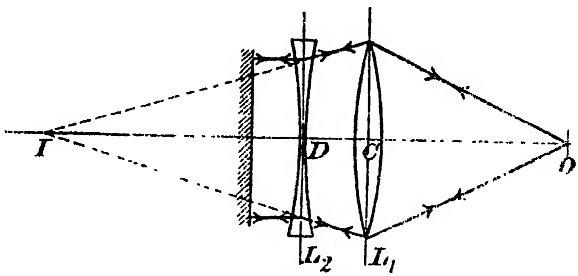


FIG. 22·8.

is produced alongside the object. This is only possible when all the rays from O after refraction through the two lenses fall normally on the plane mirror, i.e. the rays refracted through the concave lens are parallel to the axis CD. The distance DI is therefore numerically equal to the focal length of the lens L_2.

The Refractive Index of the Material of a Lens.—We have already learnt [cf. p. 424] that the focal length of a lens is related to the radii of curvature of its two faces by the formula

$$\frac{1}{f} = (\mu - 1)\left[\frac{1}{r_1} - \frac{1}{r_2}\right]$$

where the symbols have their usual meanings. If the lens used is a double concave one the radii of curvature of its two faces may be found at once by the method indicated on p. 436. When one or both of the surfaces are convex, however, the method is not so direct. Let L, Fig. 22·9 (*a*), be a double convex lens, the centres of curvature of its faces being C_1 and C_2, while r_1 and r_2 are the

corresponding radii of curvature respectively. Consider a ray OA from a point object O so placed that the refracted ray travels along AB, the normal to the second surface of the lens at B. At this point the energy of the incident light is partly reflected and partly transmitted. [This happens in all cases when light is incident upon a boundary separating two media, unless the angle of incidence is greater than the critical angle concerned.] In the present instance, however, the reflected portion returns along the path BAO and is used in setting the lens in the desired position with respect to O. The transmitted portion of the energy appears to come from C_2. These statements apply to all paraxial rays from O when, after refraction at the first surface of the lens, they fall normally upon the second surface. If therefore one looks

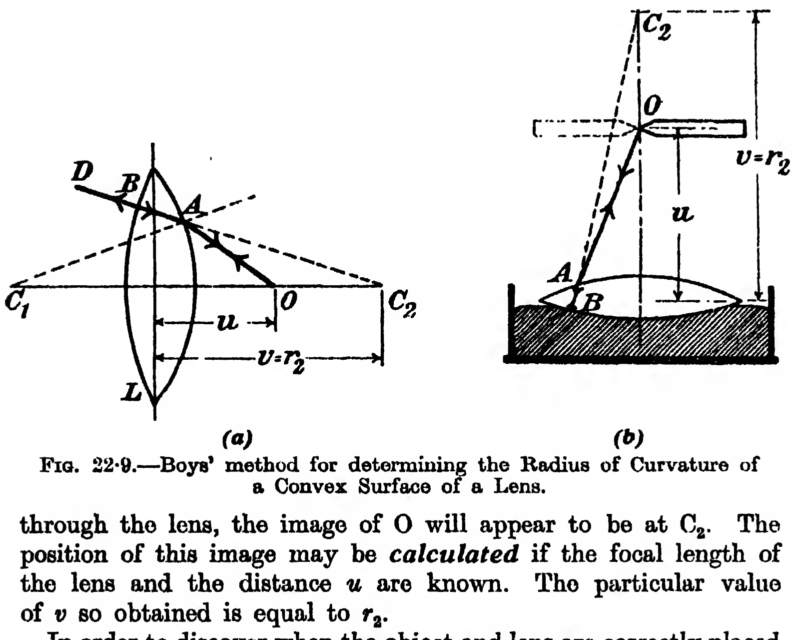


FIG. 22·9.—Boys' method for determining the Radius of Curvature of a Convex Surface of a Lens.

through the lens, the image of O will appear to be at C_2. The position of this image may be ***calculated*** if the focal length of the lens and the distance u are known. The particular value of v so obtained is equal to r_2.

In order to discover when the object and lens are correctly placed for the above conditions to hold, use is made of that portion of the light energy which is ***reflected*** from B along BA and after refraction at A produces an image at O. The distance u is then measured and r_2 calculated.

By reversing the lens the radius of curvature of the other surface may similarly be found. The value of μ for the material of the lens may then be calculated.

The above experiment is most easily carried out with the aid of illuminated cross-wires, etc., but it may also be performed with

a pin as object if the lens is floated on clean mercury. The pin is arranged horizontally and its position adjusted until there is no parallax between the pin and its image—cf. Fig. 22·9 (*b*). In this particular instance nearly all the light energy is reflected at B, but the theory is as before.

The Refractive Index of a Liquid.—For this experiment we require a converging lens. Its focal length is first determined using a pin and plane mirror. For convenience the lens is placed in contact with the mirror. A small quantity of the liquid is then placed between the lens and mirror, forming a liquid plano-concave lens. The focal length of the combination is then determined. If this is f, while f_1 and f_2 are the focal lengths of the glass and liquid lenses respectively, $\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$, so that f_2 may be deduced. Now the radius of curvature of the concave surface of the liquid lens is equal to r_2 the radius of curvature of that surface of the convex lens in contact with it. If this is known the refractive index of the liquid may be calculated, for

$$\frac{1}{f_2} = (\mu - 1)\left[\frac{1}{r_2} - \frac{1}{\infty}\right].$$

The Focal Length of a Converging Lens.—A variation of the usual u and v method for the determination of this quantity is as follows :— An object OA, Fig. 22·10, is placed in front of a lens so that a real image is produced. Instead of receiving this image on a screen, however, a plane mirror is placed behind the lens and it is adjusted until an image equal in size to OA and *erect* is produced in the plane containing OA. This occurs when the real image IB is formed at the surface of the mirror so that u and v are at once known.

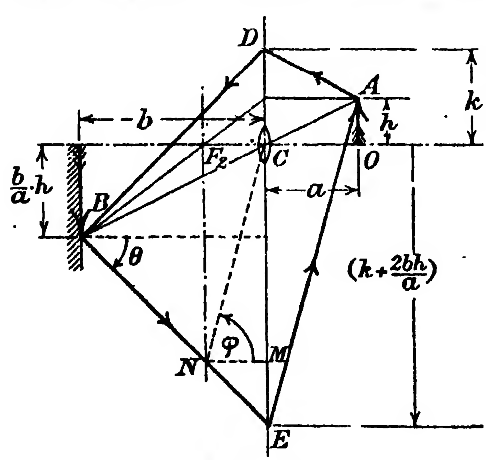


FIG. 22·10.

The 'coincidence' of image and object is so perfect in this instance that even if the mirror is slightly tilted the coincidence is not destroyed. If a pin is used as object there is no difficulty in determining when the image and object do coincide, but if illuminated cross-wires are used it is preferable to arrange a thin sheet of glass at 45° to the axis and receive the image on a piece of ground glass at the same distance from the glass as is the object.

To show that the image fulfils all the conditions stipulated above, let us consider any ray AD which after refraction through the lens travels along DB. At B it is reflected along BE, the angles of incidence and reflexion being equal. To determine the path of the ray BE after refraction through the lens we may use a graphical construction as follows:—If BE cuts the second focal plane of the lens in N, the refracted ray EA will be parallel to the secondary axis NC.

The validity of the above construction may be established as follows. If h is the length of the object, a and b the distances of the object and image from the lens centre [all these quantities being mere numbers], the length of the image is $\frac{b}{a}\cdot h$, while the distance CF_2 is $\frac{ab}{a+b}$. If k is equal to CD, then $CE = k + 2\frac{b}{a}h$. If θ is the angle indicated,

$$\tan\theta = \left(\frac{k + \frac{b}{a}\cdot h}{b}\right),$$

so that

$$EM = \left(\frac{ab}{a+b}\right)\tan\theta = \frac{(ak+bh)}{a+b}.$$

$$\text{Hence } \tan\phi = \frac{\left(k + 2\frac{b}{a}\cdot h\right) - \frac{(ak+bh)}{(a+b)}}{\left(\frac{ab}{a+b}\right)}.$$

Let the ray through E parallel to NC cut OA in A_1. Then the vertical distance between A_1 and E is

$$a\tan\phi = \frac{a\left[k + 2\frac{b}{a}h\right] - \frac{a(ak+bh)}{(a+b)}}{\left[\frac{ab}{a+b}\right]}$$

$$= \frac{1}{b}\left[k + 2\frac{b}{a}h\right](a+b) - \frac{1}{b}(ak+bh)$$

$$= h + k + 2\frac{b}{a}h.$$

But this is the vertical distance between A and E, so that A and A_1 must coincide. Since AD was any arbitrary ray it follows that all rays from A return to A after reflexion at the mirror.

Similarly, rays from any point in OA return to the same point after reflexion at the mirror.

Experimental Determination of $_g\mu_w$.—Light from an arc lamp (not shown) is concentrated on an aperture, S, Fig. 22·11. L_1 is a converging lens arranged so that S is at its first principal focus. The beam of parallel light emerging from L_1 falls on a plane mirror, M, and is reflected downwards into a deep vessel, A, containing water coloured with fuschine. A plano-convex lens, L_2, is supported as shown, its plane surface being uppermost. In this way refraction takes place at the curved glass-water interface. The light is brought to a focus at F_2. Let v be the distance of F_2 below the pole of

the curved surface of L_2. Let μ_1 and μ_2 be the refractive indices of glass and water respectively, so that ${}_g\mu_w = \frac{\mu_2}{\mu_1}$. Applying the formula

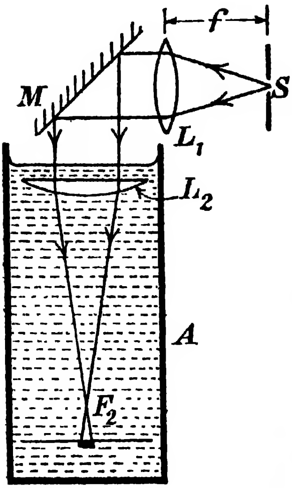


FIG. 22·11.—Apparatus for determining ${}_g\mu_w$.

$$\mu_1\left[\frac{1}{r} - \frac{1}{u}\right] = \mu_2\left[\frac{1}{r} - \frac{1}{v}\right],$$

and, remembering that $u = \infty$, we have

$$\frac{\mu_1}{r} = \mu_2\left[\frac{1}{r} - \frac{1}{v}\right].$$

Hence, if r is known, $\frac{\mu_2}{\mu_1}$, i.e. ${}_g\mu_w$ may be found.

[If fuschine is not available the focal point F_2 may be located by allowing a piece of white cardboard, loaded on its underside with lead shot, to fall through the liquid. The position of F_2 is indicated by a bright spot of light which appears on the cardboard as it passes downwards through F_2—see the diagram.]

Example.—A hemispherical bowl with a radius of curvature of 8·5 cm. contained carbon bisulphide and was placed so that its curved surface was in contact with water in a tall jar. Parallel light was directed vertically downwards and brought to a focus at a point 42 cm. from the curved surface. Compare the refractive indices of the two liquids.

Let μ_1 and μ_2 be the refractive indices of carbon bisulphide and of water respectively. In the equation

$$\mu_1\left[\frac{1}{r} - \frac{1}{u}\right] = \mu_2\left[\frac{1}{r} - \frac{1}{v}\right],$$

we must write $u = \infty$. Then, remembering that r is positive and v negative, we have

$$\mu_1\left[\frac{1}{8\cdot5} - 0\right] = \mu_2\left[\frac{1}{8\cdot5} + \frac{1}{42}\right].$$

$$\therefore \frac{\mu_1}{\mu_2} = 1\cdot20.$$

CHAPTER XXIII

SPECTRA AND RELATED PHENOMENA

The Composite Nature of White Light.—Whilst attempting some experiments with lenses he had constructed, NEWTON noticed that the images were blurred and indistinct, whereas the images produced with the aid of curved mirrors were much sharper. Newton made other lenses, taking greater care with the polishing operations, but the trouble still persisted. He surmised that the fault lay not in the lenses but perhaps in the laws of refraction or in the nature of light itself. Sir Isaac had been using sunlight for these experiments, and proceeded to make the following tests :— Sunlight was allowed to enter a darkened room through a circular hole one-third of an inch in diameter and an image of the sun was obtained on a screen 15 feet away. Then a glass prism was placed in the path of the rays so that the rays were deviated upwards. Newton noticed that there was an elongated image on the screen and that it was coloured at its extremities. Other experiments in which a narrower slit was used were then made and Newton found that the image was coloured along its whole length. He distinguished seven different colours—red, orange, yellow, green, blue, indigo, and violet. Since, however, the colours gradually pass from one to the other, the actual number of colours is really infinite. The image produced in experiments similar to the above is termed a ***spectrum.***

Newton also selected one of the above seven colours by allowing it to pass through another slit, and then placed a second prism in its path—cf. Fig. 23·1 : the ray was deviated still more, but no new colours were formed. Newton also permitted the whole of the coloured spectrum to fall on another prism having its refracting edge in a direction opposite to that of the first prism. A white image was obtained. If, however, one colour was removed from the first image the recombination of the remaining colours did not produce white light.

A piece of red glass was placed in front of the prism, when only the red portion of the spectrum was obtained ; and on replacing the

red glass by a piece which was blue only the blue end of the spectrum was observed.

From these and other experiments Newton concluded that sun-

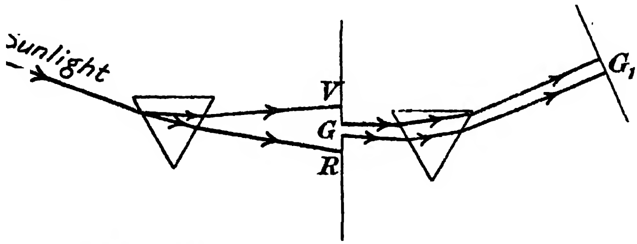


FIG. 23·1.—Newton's Experiment on the Formation of a Spectrum.

light, or white light as it is generally termed, was a mixture of several colours and that the prism merely served to separate out the constituent colours. The white light is said to have been ***dispersed.***

Newton's Experiment with Crossed Prisms.—Two prisms, ABC and LMN, Fig. 23·2, were arranged so that their refracting edges were at right angles to each other and a beam of sunlight was allowed to pass through the combination. Newton noticed that the violet and blue end of the spectrum which suffered the greatest

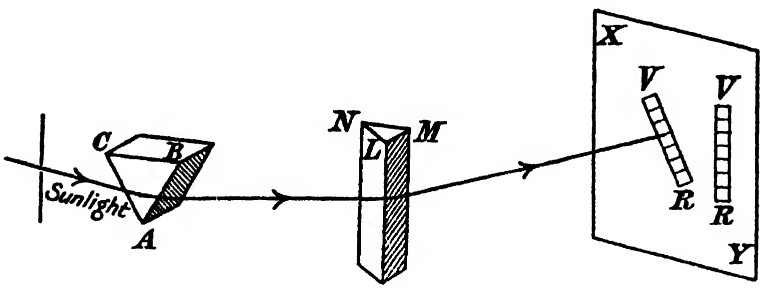


FIG. 23·2.—Newton's Experiment with Crossed Prisms.

deviation after refraction through one prism was the most deviated portion after refraction through the second prism so that when the spectrum was received on a screen XY it was found to be inclined to the vertical although, in the absence of the second prism, a vertical image was obtained.

Dispersion.—The formation of a spectrum by the means previously indicated is due to the fact that the constituent colours present in the white light have been separated by the material of the prism. This is said to have produced ***dispersion*** of the

heterogeneous light incident upon it. By using prisms of different materials it is easy to show that the dispersion depends upon the material.

If the source of light is ***monochromatic*** the spectrum consists of a single line, the prism merely deviating the rays of light passing through it. By using such a source and different prisms, each in the position of minimum deviation for the particular light used, it is found that the deviation depends upon the material of the prism. Now it is never found that a prism set in the position of minimum deviation for one set of rays is in the position of minimum deviation for all colours. In practice, when the visible region of the spectrum is being explored, the prism is set in the position of minimum deviation for yellow rays (from sodium).

The Projection of a Spectrum on a Screen.—A converging lens L_1, Fig. 23·3, focuses the image of a powerful source of light O on a slit S. Another converging lens L_2 is used to produce an image of the slit at S_1. The prism ABC, with its refracting edge parallel to the length of the slit, is then introduced when, instead of the image at S_1, there appears a coloured band somewhere in the neighbourhood VR. The definition of the image will be considerably improved by rotating the prism until it is in the position of minimum

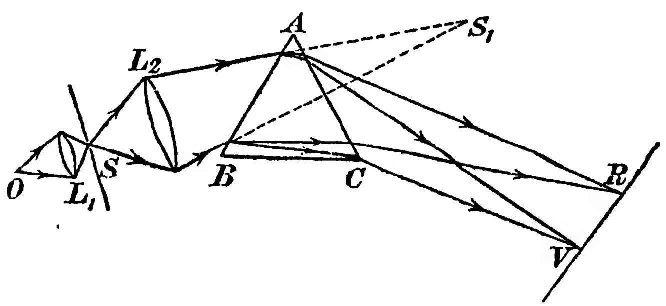


FIG. 23·3.—Projection of a Spectrum on a Screen.

deviation. It will also be improved by narrowing the slit and by adjusting L_2 slightly. The spectrum obtained will be moderately pure.

In connexion with this experiment it is important to realize that the shape of the slit exercises a marked influence on the purity of the spectrum. It must be remembered that the optical arrangement really produces a series of images of the slit—one for each constituent colour. If, therefore, the slit is wide, the different images may overlap and the spectrum will not be pure. If, for example, the slit is replaced by a circular hole, there is produced a series of coloured images of the hole on the screen. Only the

outer edge of the complete image will be coloured—red at one end, blue at the other—the central region being white, i.e. the colours have here recombined to produce white light. Similarly, if the slit is replaced by another of any other form, the individual parts of the spectrum will assume the new shape of the slit.

The Pure Spectrum.—If neighbouring colours in a spectrum do not overlap, the spectrum is said to be pure. The spectrum formed with the above arrangement of prism, lenses, slit, etc., is only tolerably pure, i.e. its purity is only partial since the rays of any one colour passing through the prism are not parallel to one another, and therefore the prism cannot be set so that they all pass through it and suffer a minimum deviation—a necessary condition for the formation of a pure spectrum when the light incident on the prism is not parallel. It is more usual for the incident light to be parallel, a condition obtained by placing the slit S in the first focal plane of a converging lens L_2—cf. Fig. 23·4. If the

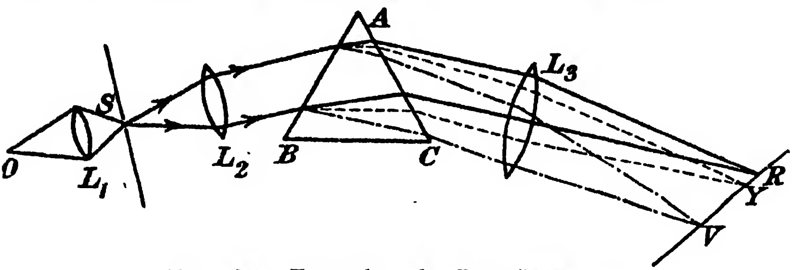


FIG. 23·4.—Formation of a Pure Spectrum.

source is not monochromatic each colour present gives rise to a parallel beam of light within the prism and also after refraction through the prism. A third converging lens L_3 collects the emergent parallel beams and brings each to a focus so that a brilliant image VR is obtained. If measurements are to be made the prism is usually adjusted so that it is in the position of minimum deviation for sodium (yellow) light, but this does not affect the purity of the spectrum.

The spectrum may be produced on a screen, i.e. observed objectively, or examined by a converging lens of suitable focal length placed so that a virtual magnified image of the spectrum is formed—cf. p. 500. The spectrum is then said to be examined subjectively. It should be noted that VR is not normal to the axis of L_3: this is because the focal length of the lens is shorter for violet rays than for red. If, however, L_3 is an achromatic lens [cf. p. 463], VR is normal to the axis of the lens.

The Spectrometer.—One of the simple forms of this instrument which was designed to examine various spectra is shown in Fig. 23·5.

A prism P is rigidly fixed on a table which is capable of rotation about an axis passing through the centre of a divided circle D. A narrow slit S, whose width can be adjusted, is illuminated by a source of light and is caused to lie at the focus of an achromatic [1] converging lens L_1, so that a beam of parallel light is incident upon a face of the prism—the slit and lens L_1, together with the tube in which they are mounted, constitute the ***collimator***. The distance from the source to S must be such that the spectrometer is not damaged by becoming too hot, but that the whole of the lens L_1 is filled with light. After refraction through the prism the light consists of several beams of various colours unless the incident beam is monochromatic; this refracted light enters a telescope T,

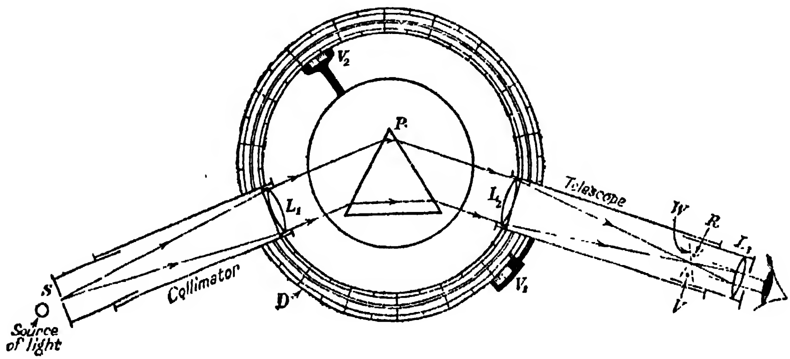


FIG. 23·5.—A Spectrometer.

and each constituent beam is focused by means of L_2, the front lens of the telescope; this lens is similar to L_1. The eye-piece L_3 produces a magnified image of the spectrum [cf. chap. XXV], and enables the colours to be seen clearly. The image (spectrum) is seen at infinity if the eye is at rest. In the diagram the prism has been drawn in the position of minimum deviation for sodium light and only the paths of the yellow rays are shown: the points to which red and violet rays would be focused are indicated, however.

It will be noticed that because the position of the focal plane of an achromatic lens is independent of the wave-length of the light used, the spectrum VR is normal to the axis of the telescope.

Verniers, V_1 and V_2, rigidly attached to the telescope and table respectively, enable the settings of these parts of the complete instrument to be ascertained with accuracy. [In actual practice on all reliable instruments the number of verniers is doubled so

[1] An achromatic lens is a double lens having a focal length practically independent of the colour of the light used [cf. p. 462].

that errors of centring may be eliminated by reading each pair of verniers: these pairs are not indicated on the diagram.]

The Adjustments of a Spectrometer.—The eye-piece is moved backwards or forwards until the cross-wires, which are placed in the observing telescope T, are visible. A distant object is then viewed through the telescope, and the tube carrying the eyepiece and cross-wires is moved until a clear image of the distant object coincides with the wires—the coincidence is verified by moving the eye sideways; if the image does not apparently move with respect to the cross-wires, then the adjustment has been accomplished. After this adjustment the eye-piece must not be disturbed. The slit of the collimator is then illuminated and the telescope directed towards it; the slit is moved to and fro until a clear image is formed on the cross-wires. The telescope having been adjusted so that parallel rays are brought to a focus on the cross-wires, it follows that the slit is at the focus of L_1, so that a parallel beam of light is incident upon the prism. The slit must be narrowed to permit accurate measurements to be made. If it is necessary to level the prism the following procedure will be effective:—

The term 'levelling the prism' is somewhat ambiguous; what is really implied is that the refracting edge of the prism must be made parallel to the axis of rotation of the table. Three screws A, B, C, Fig. 23·6, enable the table to be raised or lowered. The prism is placed with its edge *ab* perpendicular to an imaginary line joining the two screws B and C. The collimator is placed so that both the faces *ab* and *ac* receive light. When the telescope receives light reflected from *ab* the screws B and C are adjusted till the image is in the centre of the field. The telescope is then turned to receive light from *ac* and the screw A adjusted until the image is again in the centre of the field. The telescope is then moved to its former position, and B adjusted; the telescope is then brought in front of *ac* and A adjusted; the process is repeated until the image is always in the centre of the field irrespective of the face from which the light is reflected.

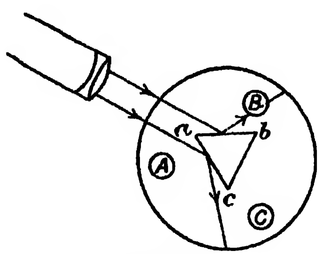


Fig. 23·6.

When no distant object is available the following method due to Schuster is adopted:—The prism having been placed on the table so that it is in line with the collimator and telescope, the telescope is rotated until a blurred image is seen in the telescope. The table is then slowly moved and at the same time the telescope rotated, so that it always receives the image. If the table is being

rotated in the appropriate direction the image will soon pass through the position of minimum deviation. If the telescope is moved back through a few degrees it follows that it will receive light and form an image for two positions of the prism—A and B [dotted], Fig. 23·7. The prism is rotated until it is in the position B, i.e. the light falls more obliquely on the prism than when it occupies the position of minimum deviation. The telescope is then focused

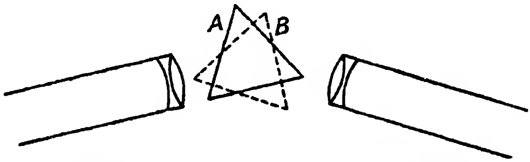


FIG. 23·7.—Schuster's Method of Adjusting a Spectrometer.

until the image is sharp and without parallax on the cross-wires. The prism being rotated to the position A the image will be blurred, but it may be focused again by moving the slit of the collimator. The prism is then put in the position B and a sharp image produced by adjusting the eye-piece of the telescope. The process is continued until the image is always sharp and without parallax on the cross-wires. When this has been achieved the collimator is producing a parallel beam of light which is brought to a focus on the cross-wire of the telescope.

To understand the principle here involved let us consider Fig. 23·8. Suppose that A, B, and C are three different positions of a converg-

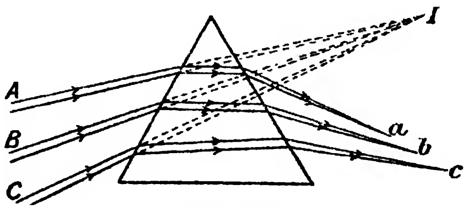


FIG. 23·8.

ing beam of light incident upon a prism such that the angles of incidence are respectively less than, equal to, and greater than that corresponding to the position of minimum deviation. Suppose *a*, *b*, and *c* to be the images produced. The less oblique pencil A becomes less parallel after refraction, while the more oblique one becomes more parallel. When the prism is in the position B, Fig. 23·7, the rays are as at C, Fig. 23·8. They will therefore enter the telescope in a more parallel position than when the telescope

was in the position of minimum deviation. It is therefore necessary to adjust the telescope to improve the definition of the image. When the prism is in the position A, Fig. 23·7, the light entering the telescope will be too convergent, but this may be rendered more parallel by adjusting the collimator.

To Determine the Optical Constants of a Prism.—The spectrometer furnishes a very accurate means of determining the refractive index of the material of a prism by measuring the angle of the prism and then the angle of minimum deviation. It has been shown that the deviation produced by a prism depends upon the colour of the light; hence monochromatic light, i.e. light of a single colour, must be used. The support for an upright gas mantle is heated to redness and dipped into a mixture of borax and sodium chloride; the mixture is then fused in a bunsen burner,

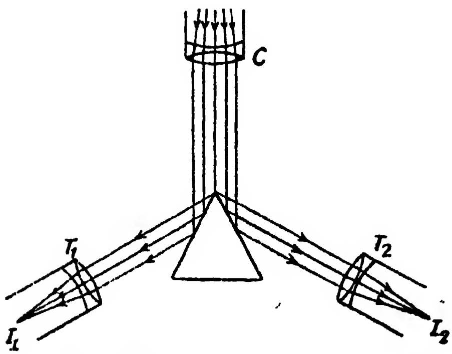


Fig. 23·9.—Measurement of the Refracting Angle of a Prism.

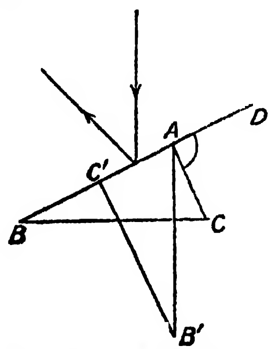


Fig. 23·10.—Measurement of the Refracting Angle of a Prism.

when the whole flame is coloured yellow. This forms a convenient and almost monochromatic illuminant for the slit of the spectrometer.

(*a*) ***The Angle of the Prism.***—The prism is arranged so that light from the collimator C is incident upon both faces at the same time [Fig. 23·9]. The image I_1 is focused on the cross-wires of the telescope, the position of the telescope being observed. The telescope is then moved to receive the image I_2 on its cross-wires, its position again being observed. The angle through which the telescope has been rotated is twice the angle of the prism [cf. p. 410].

A second method of determining the angle of the prism will be understood from Fig. 23·10. When the image from one face AB has been sighted in the telescope, the latter is kept in that position and the prism rotated until the edge AC occupies a position coincident with or parallel to AB. For this to occur the prism will have

to be rotated through an $\widehat{DAC}$. It is at once apparent that $\pi - \widehat{DAC}$ is equal to the angle of the prism.

(*b*) ***The Angle of Minimum Deviation.***—The image of the slit, formed by refraction, is observed in the telescope and the prism is rotated, the image being followed with the telescope, until the image appears to be stationary—the prism is then in the position of minimum deviation for rays of the particular colour characteristic of the sodium flame. A further rotation of the prism and the image recedes. The prism is then rotated and the position of minimum deviation on the other side of the 'line of fire' found. Half the angle through which the telescope has been rotated is the angle of minimum deviation γ.

Then $$\mu = \frac{\sin \frac{1}{2}(\alpha + \gamma)}{\sin \frac{1}{2}\alpha},$$

where α is the angle of the prism.

Spectrum Analysis.—(i) ***Line Spectra.*** If, whilst the prism is on the spectrometer table, the sodium flame is removed and a bunsen flame [fed with various salts, such as those of lithium, potassium, etc.] substituted, then coloured images of the slit will be observed in the telescope. Each image corresponds to a definite colour in the light emitted from the ***vapour*** of the substance—each image is referred to as a line in the spectrum.

By certain means, some of which are discussed later, it is possible to determine the wave-length of the light corresponding to each line in the spectrum. If, for example, a sodium salt is heated in a bunsen flame the latter assumes a brilliant yellow tinge and the spectrum, in the visible region, consists of two bright lines in the yellow region. They are very close to each other and together constitute the D_1 and D_2 lines of the spectrum. Their wave-lengths are respectively 5896 and 5890 Ångstrom units, where one Ångstrom unit—written 1 Å—is equal to 10^{-8} cm.

Spectra of the different gases are obtained when electric discharges pass through a tube containing the particular gas in a rarefied condition and the light examined spectroscopically; also when an electric arc is produced between metal electrodes or an electric spark is passed between them, the light in each instance being examined with a spectroscope.

Bunsen and Kirchoff were the original investigators in this wide field of research known as ***spectrum analysis.*** One of their greatest discoveries was that ***under given conditions each element emitted light producing a definite spectrum which was characteristic of that element only.*** Having made this discovery they began a systematic examination of the light from certain important stars, and from the lines in the spectra to which

they gave rise they established with certainty the existence of several elements present in those stars. Later on, a spectroscopic examination of the light from the sun revealed the fact that there was present in the sun an element then unknown on the earth. It was called ***helium***—a gas now known to be present in the atmosphere.

During the present century a detailed examination of the spectra of the elements has enabled us to gain an insight into the structure of matter.

From the above remarks it will appear that an examination of the spectrum of the light emitted by a certain substance when in a gaseous condition enables us to detect its composition qualitatively; in recent years, by paying particular attention to certain lines in spectra, it has become possible to estimate the actual percentage composition of a mixture.

Usually a spectrum of the light from the arc produced between metal electrodes is a line spectrum characteristic of that metal—it is termed an ***arc spectrum*** and it is now known that the corresponding colours are due to radiation emitted by an atom.

When a ***spark spectrum,*** i.e. the spectrum given by the light from a spark between metal electrodes, is examined, the lines are not in the same positions as when the light from an arc between the same electrodes is observed. The spark spectrum is produced by the radiation from atoms which have lost one electron, i.e. from ionized atoms, or atoms carrying a positive charge of electricity.

The Continuous Spectrum.—If the light from a white-hot body, such as the filament of an electric lamp, is examined with the aid of a spectroscope, the spectrum is found to be continuous, i.e. all the lines in the visible region appear in the spectrum of the light from such a body. The reason for this is that the atoms get into a state where they are able to emit radiation, but cannot emit it as when they exist as a vapour. Light of all wave-lengths is therefore emitted and the spectrum appears continuous.

Band Spectra.—In an emission spectrum the lines sometimes appear crowded together in certain regions. When they are examined with the aid of an ordinary spectroscope they appear as continuous flutings or bands. The line structure of such a band is only revealed with spectroscopes possessing a high resolving power, i.e. they are able to separate lines differing by only a small fraction of an Ångstrom unit. We now know, that whereas line spectra are due to atoms, band spectra are due to molecules. Such spectra appear very frequently in the infra-red region [cf. p. 472].

Absorption Spectra.—When the light from the tungsten filament of an electric globe or other source of white light is focused

on the slit of a spectrometer a continuous spectrum is seen in the telescope. If a piece of red glass is placed before the slit the image is seen to be red. The glass has robbed the white light of some of its constituents so that only the remaining colours are present and it is only these which are analysed by the spectroscope. The red image now obtained is referred to as the absorption spectrum of red glass. The absorption spectrum of a vapour, which when hot emits bright lines, is found to consist of a continuous spectrum crossed by dark lines in exactly those positions previously occupied by the bright lines of the emission spectrum. This shows that the cold vapour absorbs mainly those colours which it emits when its temperature is sufficiently high. This point may be demonstrated further by the following experiment:—The spectrum of an electric arc is projected on a screen—cf. the apparatus described on p. 450. A bunsen flame tinged with sodium is placed behind the illuminated slit—the spectrum is crossed by a dark line in the yellow region—its position corresponds to that of the so-called D-line of the sodium spectrum [really there are two D-lines very close together, but unless a large prism and a narrow slit are used they will overlap and appear as one—we say that the prism has failed to resolve the sodium lines]. On introducing some sodium salt into the electric arc a bright yellow line appears in the place of the dark one originally present. This proves that the sodium flame absorbs the same waves as it emits. The bright sodium lines appear if an opaque screen is placed before the source of white light.

The absorption spectrum of an aqueous solution of potassium permanganate may be investigated as follows:—The spectrum of the light from an arc lamp is produced on a screen and a glass cell containing water is then placed at some convenient position in the path of the beam. By dipping a glass rod into a strong solution of the permanganate and rinsing it in the water the concentration of the latter may be gradually changed by repeating the process. The appearance of at least two dark bands, one in the green and one in the blue, is soon noticed. As the concentration increases the two bands widen out and finally coalesce. Finally, the whole of the visible spectrum disappears showing that a strong aqueous solution of potassium permanganate is opaque to all visible radiations.

The absorption spectrum of a dilute solution of didymium chloride is very interesting since it contains five or six bands. The solution may be contained in a small weighing bottle and used to concentrate an image of the filament of an electric lamp (filament vertical) on the slit of the spectroscope. The spectrum is then examined in the usual way.

These experiments are more likely to succeed if a hollow prism

filled with carbon bisulphide is used since the dispersion is greater than with a glass prism.

The Solar Spectrum.—By reflecting a beam of sunlight on to the slit of a spectrometer its spectrum will be found to consist of a continuous background crossed by many dark lines. These are termed the FRAUNHOFER lines in honour of the man who discovered them. BUNSEN and KIRCHOFF showed that they were in the same positions as some of the bright lines of the emission spectra belonging to terrestrial elements; they concluded that these elements were present in the sun. It is now known that the dark lines are due to the absorption of some of the white light emitted by the central and hotter regions of the sun as it passes through the cooler envelope [the ***chromosphere***] surrounding the sun.

The Recombination of Spectral Colours.—Reference has already been made [cf. p. 447] to the fact that NEWTON used a second prism, with its refracting edge pointing in an opposite direction to that of the first prism, in order to show that the colours of the complete spectral fan could be recombined to produce

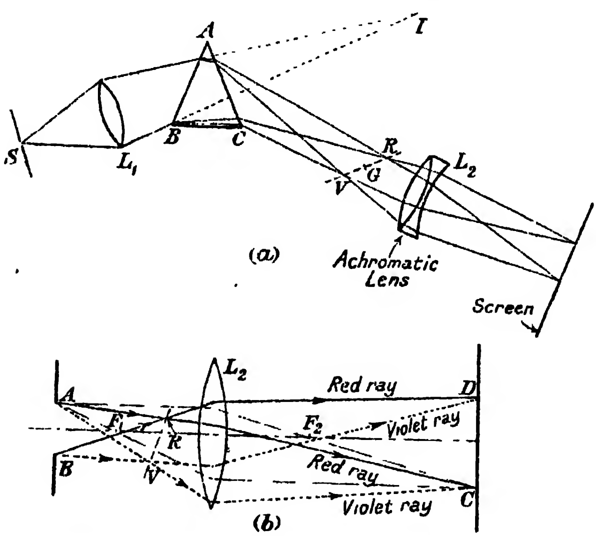


FIG. 23·11.—Recombination of Spectral Colours to form White Light.

white light. Another arrangement—known as a colour patch apparatus—whereby the same and other facts may be established is as follows. A slit, S, Fig. 23·11 (*a*), is illuminated by the light from an arc lamp, and a converging lens L_1 arranged so that in the absence of the prism ABC a real image of the slit is formed at I. The prism, for preference a hollow one filled with carbon

bisulphide, is then placed in position so that a pure spectrum is produced at RV. L_2 is a converging lens arranged so that it gives rise to an image of the face AB of the prism on a distant screen. If the adjustments have been carried out properly this final image will appear white.

To understand the formation of this white image let AB, Fig. 23·11 (*b*), be an aperture in a screen—this corresponds to the face AB of the prism. L_2 is a converging lens producing an inverted image CD of AB. The construction lines are shown thin. Now suppose that ARB and AVB are beams of red and violet light respectively—in the actual arrangement at (*a*) these are produced after the light has passed through the prism, but we may assume that they are formed by rays from two different sources on the left of the diagram (*b*) to simplify the drawing. Since C is the image of A it follows that the ray AR must travel to C after refraction through L_2: similarly BR, a red ray from the bottom of the aperture, travels to D. Thus CD is covered with red light. Similarly it is covered with violet light—in fact, with light of all the other spectral colours. These superimposed colours produce white light.

Colour Mixture and Complementary Colours.—If a screen with a slit parallel to that at S is placed along RV then a pure spectral colour will be produced at CD. If, however, a pencil is placed so that it intercepts the bluish-green rays, i.e. the pencil is at G, then the resulting colour on the screen will be due to all the other rays in the spectrum. This colour is red. Similarly if the blue rays are intercepted the colour will be orange. The colours produced on the screen under these conditions are termed ***mixed colours.***

The above experiments show that a complex light stimulus cannot be analysed into its spectral components by the human eye. In this respect, the eye is essentially different from the ear which is able to analyse a musical note into its components, i.e. to distinguish between one compound note and another. Moreover, in acoustics, like sensations are produced by like causes—but this is not true in connexion with light sensations. For example, an aqueous solution of potassium dichromate appears to possess a deep orange colour when viewed by transmitted light. As Abney first showed, if this light is examined spectroscopically it is found to consist of several components—in fact, there is produced a spectrum from which the blue and violet rays have been removed completely and the green rays in part. On the other hand, it is possible to select a small region of the spectrum VR such that the colour of the patch on CD is exactly the same as that of the light transmitted by the bichromate solution.

Primary Colours.—Abney defines a primary colour as one which cannot be formed by the mixture of two or more colours. The original investigators in colour phenomena were the artists, who found that the pigments on their palette were not capable of being mixed to produce red, yellow or blue. They termed these the primary colours. When the physicist became interested in this subject he discovered that the yellow was not a primary colour since it could be obtained as a mixture of red and green. On the other hand, green was shown to be a primary colour. These facts are easily verified with the apparatus shown in Fig. 23·11 (*a*). It is only necessary, for example, to place slits along VR so that the red and green rays are transmitted and the colour of the patch at CD is yellow.

Modern work has shown that violet, and not blue, is a primary colour.

Complementary Colours.—When the sensation of white light is produced by two colours, those colours are said to be complementary to each other. The colour patch apparatus described above may be used to obtain the complementary colours. One of the colours produced at VR is cut out by means of an opaque rod, and the remaining colours combined by the lens L_2 to form an image at CD. The resultant colour is complementary to that which has been removed. Since the three primary colours, taken in the correct proportions, produce white light, it follows that if one of these is removed, the combination of the other two will give the corresponding complementary colour.

TABLE OF SOME COMPLEMENTARY COLOURS

Red	Orange	Yellow	Green	Bluish-green	Blue	Violet	Colour blocked out by rod
Bluish green	Blue	Indigo	Purple	Red	Orange	Greenish yellow	Resulting colour mixture

The colours in the lower line are complementary to the corresponding colour in the upper line. The remarks in the last column indicate how the existence of these colours may be verified with the colour patch apparatus.

In connexion with complementary colours it must be emphasized that the white light arising from the combination of two pure complementary colours is not physically identical with sunlight; the impression of white due to the superposition of two complementary colours is due to a peculiarity of the eye [cf. p. 495],

and if such light is subjected to analysis by a spectroscope, the complete spectral fan is not obtained. The colours then seen are, of course, the two complementary colours.

Angular Dispersion and Dispersive Power.—We have seen that the solar spectrum is crossed by a number of dark lines, parallel to the length of the slit. The colours corresponding to the more prominent ones are denoted by the letters A, B, C, D, E, F, G, H. The first three are in the red, D in the yellow, E in the green, F and G in the blue, and H in the violet region of the spectrum. The lines C and F are due to hydrogen and may be produced by connecting a glass tube containing this gas [or water vapour] at a pressure of less than 1 mm. of mercury, to the secondary terminals of an induction coil. If water vapour is used the hydrogen lines appear, because the electric discharge breaks up the water molecule.

When a parallel beam of white light passes through a prism, rays of different colour are inclined to one another, although for any one colour the rays are all parallel. For any two colours the angle between them is termed the ***angular separation*** caused by the prism for these two colours and for the particular angle of incidence involved. If μ_b and μ_r (or μ_F and μ_C) are the refractive indices of a material for blue and red rays, and μ (or μ_D) is the index for sodium light, the quantity $\dfrac{\mu_b - \mu_r}{\mu - 1}$, $\left(\text{or } \dfrac{\mu_F - \mu_C}{\mu_D - 1}\right)$, is termed the ***dispersive power***, ω, of the material with reference to the three wave-lengths chosen. When the material is in the form of a prism with a small refracting angle, it can be shown that the ratio of the angular separation for two rays of different colours to the deviation of the mean ray is equal to the *dispersive power* of the material. For in general, if μ is the refractive index of the material and ψ the deviation for a prism whose refracting angle α is small, then, with the notation used on p. 409, we have

$$\psi = (i_1 - r_1) + (i_2 - r_2) = (\mu - 1)\,(r_1 + r_2),$$

for $i_1 = \mu r_1$ and $i_2 = \mu r_2$ when the angles of incidence, etc., are small as we now assume—cf. p. 413. But $r_1 + r_2 = \alpha$, so that $\psi = (\mu - 1)\alpha$. If μ_D is the refractive index of the material for sodium light,

$$\psi_b = (\mu_b - 1)\alpha = \frac{\mu_b - 1}{\mu_D - 1}.\,(\mu_D - 1)\alpha = \frac{(\mu_b - 1)}{\mu_D - 1}\cdot\psi_D.$$

Similarly the expression for ψ_r may be obtained. From these two expressions we have

$$\psi_b - \psi_r = \frac{(\mu_b - \mu_r)}{\mu_D - 1}\psi_D,$$

which establishes the theorem.

Expressed in the notation of the calculus, the dispersive power ω is given by $\omega = \dfrac{\delta\mu}{\mu - 1}$, where $\delta\mu$ is the change in refractive index for the material in passing from one colour to a neighbouring one, and μ is now written instead of μ_D. It must be observed, however, that the numerator in the above equation is not an infinitesimal $\delta\mu$ but a small finite quantity $\Delta\mu$. Thus ω, as measured, is never the actual dispersive power for a particular colour (or wave-length), but is the mean dispersive power over a range of wave-lengths. The use of the symbol $\delta\mu$ is justified in practice since it seldom exceeds two per cent. of the value of μ in the visible region of the spectrum.

Chromatic Aberration, Achromatism (or the Removal of the Primary Spectrum).—Since μ depends upon the colour or wave-length of the light used in determining it, it follows that the behaviour of an optical instrument, which is generally expressed by a formula involving μ, will depend on the nature of the light with which it is used. Thus, when a converging lens is used to obtain a real image on a screen the image is tinged blue when the screen is held in the position B, Fig. 23·12 (a), and red in the position R. This is because the blue rays are more deviated than the red. This shows that the focal length of a convex lens is less for blue

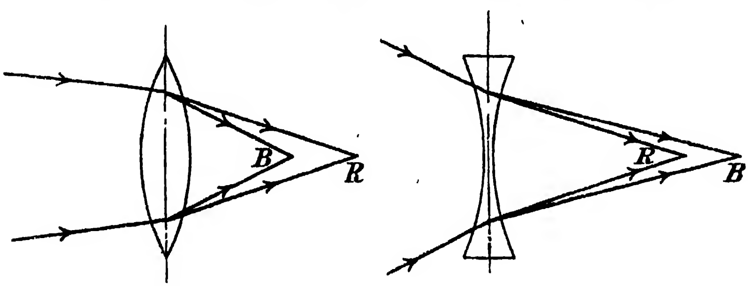


FIG. 23·12.—Chromatic Aberration of Lenses.

rays than it is for red. All the colours in the image thus produced are termed a ***primary spectrum.*** Similarly, if convergent white light falls on a diverging lens such that a real image is produced, Fig. 23·12 (b) the red image is nearer to the lens than is the blue, i.e. the focal length of a concave lens is also less for blue light than for red. This suggests that if a convex and a concave lens are combined it may be possible to obtain an image free or nearly free from colour. Such a combination is said to be ***achromatic.*** Thus a convex and a concave lens of the same material, and having focal lengths which are numerically equal, would show no disper-

sion of white light. But such a combination would form a slab of glass with concentric faces and would therefore not be a lens. To obtain an achromatic lens it is necessary to use convex and concave lenses of different materials, e.g. crown and flint glass.

In an ideal optical instrument the final image would be entirely free from colour, i.e. all the coloured images formed by rays of different wave-lengths would be equal and in the same position. Usually opticians are content if two of the coloured images can be superposed. The particular colours to be combined will depend on the use for which the lens is designed. The colours still remaining in the image formed by an achromatic combination corrected for two colours constitute a ***secondary spectrum.***

Achromatic Combination of Two Thin Lenses in Contact.—The focal length of a lens for monochromatic light is determined by the equation

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

where the symbols have their usual meanings. Hence for blue and red rays

$$\frac{1}{f_b} = (\mu_b - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = \frac{\mu_b - 1}{\mu_D - 1} \cdot (\mu_D - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

$$= \frac{\mu_b - 1}{\mu_D - 1} \cdot \frac{1}{f}$$

and $$\frac{1}{f_r} = \frac{\mu_r - 1}{\mu_D - 1} \cdot \frac{1}{f_D}.$$

$$\therefore \frac{1}{f_b} - \frac{1}{f_r} = \frac{\mu_b - \mu_r}{\mu_D - 1} \cdot \frac{1}{f_D}.$$

Now as a first approximation we may write $f_b . f_r = f_D^2$, so that

$$f_r - f_b = \frac{\mu_b - \mu_r}{\mu_D - 1} \cdot f_D = \varpi f_D,$$

where ϖ is the dispersive power of the material for blue and red rays.

For two lenses of focal lengths f' and f'' in contact, the focal length, f, of the combination is given by

$$\frac{1}{f} = \frac{1}{f'} + \frac{1}{f''}.$$

Hence $$\frac{1}{f_b} = \frac{1}{f_b'} + \frac{1}{f_b''} = \frac{\mu_b' - 1}{\mu_D' - 1} \cdot \frac{1}{f_D'} + \frac{\mu_b'' - 1}{\mu_D'' - 1} \cdot \frac{1}{f_D''}$$

where μ' and μ'' are the refractive indices of the materials of the two lenses, suffixes denoting their values for light of different colours. Similarly

$$\frac{1}{f_r} = \frac{\mu_r' - 1}{\mu_D' - 1} \cdot \frac{1}{f_D'} + \frac{\mu_r'' - 1}{\mu_D'' - 1} \cdot \frac{1}{f_D''}.$$

If the combination is to be corrected for chromatic aberration arising from these two colours, $f_r = f_b$, so that

$$\frac{\mu_b' - \mu_r'}{\mu_D' - 1} \cdot \frac{1}{f_D'} + \frac{\mu_b'' - \mu_r''}{\mu_D'' - 1} \cdot \frac{1}{f_D''} = 0,$$

or

$$\frac{\varpi'}{f_D'} + \frac{\varpi''}{f_D''} = 0.$$

Since ϖ' and ϖ'' are both positive it follows that f_D' and f_D'' must be opposite in sign.

Alternative Proof.—The formula expressing the focal length of a lens in terms of the refractive index of its material and the radii of curvature of its faces is

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right),$$

and may be written $\frac{1}{f} = (\mu - 1)C$, where C is a constant for a given lens. The change in the focal length of the lens due to a change $\delta\mu$ in the refractive index of its material is given by

$$\delta\left(\frac{1}{f}\right) = \delta\mu . C.$$

Eliminating C from this equation, we have

$$\delta\left(\frac{1}{f}\right) = \delta\mu . \frac{1}{f} . \frac{1}{(\mu - 1)}$$

$$= \frac{\omega}{f}.$$

When two lenses of focal lengths, f_1 and f_2, are in contact the focal length f of the combination is given by

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}.$$

$$\therefore \delta\left(\frac{1}{f}\right) = \delta\left(\frac{1}{f_1}\right) + \delta\left(\frac{1}{f_2}\right)$$

$$= \frac{\omega_1}{f_1} + \frac{\omega_2}{f_2},$$

where ω_1 and ω_2 are the dispersive powers of the materials of the two lenses. If the combination is to be achromatic the focal length must not change with a change in μ, i.e. $\delta\left(\frac{1}{f}\right)$ must be zero. The condition for the combination to be achromatic is therefore

$$\frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} = 0.$$

Example.—**Lenses to form a converging achromatic combination of focal length 20 cm. are to be made from crown and flint glasses having dispersive powers 0·23 and 0·37 respectively. Calculate the focal lengths of the two lenses required.**

If f_1 and f_2 are the focal lengths of the two lenses, we have

$$-\frac{1}{20}=\frac{1}{f_1}+\frac{1}{f_2},$$

while the condition for achromatism is

$$\frac{0\cdot23}{f_1}+\frac{0\cdot37}{f_2}=0.$$

Hence $f_1 = -\frac{1}{1\cdot61}f_2$, and, by substitution in the first equation $f_2 = +12\cdot2$ cm. Hence $f_1 = -\frac{12\cdot2}{1\cdot61} = -7\cdot6$ cm.

We therefore require a crown glass lens of focal length $-7\cdot6$ cm. (converging), and a flint glass lens of focal length 12·2 cm. (diverging).

Example.—If the refractive indices for sodium light for crown and flint glasses are 1·5 and 1·6 respectively and the two faces in contact have radii of curvature 15 cm., calculate the radii of curvature of the other faces of the lenses, using the data of the previous example.

If the combination is arranged so that the converging lens is nearer to the object, then, with the usual notation,

$$-\frac{1}{7\cdot6}=0\cdot5\left(\frac{1}{r_1}-\frac{1}{15}\right),\text{ or } r_1=-5\cdot1 \text{ cm.}$$

Similarly for the diverging lens,

$$+\frac{1}{12\cdot2}=0\cdot6\left(\frac{1}{15}-\frac{1}{r_2}\right),$$

where r_2 is the radius of curvature of that surface of the concave lens not in contact with the converging lens.

$$\therefore\ r_2 = -14\cdot2 \text{ cm.}$$

Achromatic Prisms.—When white light passes through a prism the light is both dispersed and deviated. By combining prisms of different glasses and therefore different dispersive powers it is possible to construct a compound prism producing deviation without much dispersion. Such a combination is said to be ***achromatic.*** This is accomplished by adjusting their refracting angles so that the dispersion due to the first is approximately counteracted by the second, i.e. the emergent ray is almost free from colour. The reason for the small colour effect remaining when two prisms have been designed so that the combination is corrected for the C and F rays, say, is that the spectra produced by different prisms are not geometrically similar, and thus the other rays are dispersed slightly. In actual practice it is better not to combine the extreme red and the extreme violet rays since these are relatively faint, but the

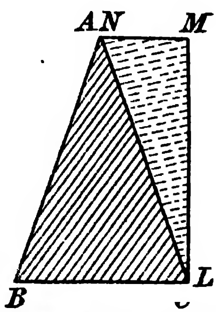


Fig. 23·13.—Achromatic Prism.

orange-yellow and blue-green, for it is such rays which exhibit the greatest difference in colour and which are also the strongest colours in the spectrum of white light.

If ABC and LMN, Fig. 23·13, are prisms of crown and flint glass respectively, with small refracting angles α' and α'', μ' and μ'' the refractive indices, ψ' and ψ'' the deviations of any ray, and if suffixes refer to colours, the angular dispersion of the first prism is

$$\psi_b' - \psi_r' = (\mu_b' - \mu_r')\alpha',$$

while that of the second is

$$\psi_b'' - \psi_r'' = (\mu_b'' - \mu_r'')\alpha''.$$

If the combination is to be achromatic for red and blue rays then

$$\psi_b' - \psi_r' = \psi_b'' - \psi_r'',$$

i.e. $(\mu_b' - \mu_r')\alpha' = (\mu_b'' - \mu_r'')\alpha''.$

This equation enables us to calculate the angle of the second prism when the first is known.

Alternative method.—It has already been shown that the deviation of a ray passing through this prism (cf. its refracting angle α is small, is given by

$$\psi = (\mu - 1)\alpha.$$

Suppose this applies to the sodium ray. Then $\Delta\psi$, the angle between the blue and red rays, is given by

$$\Delta\psi = (\mu_B - \mu_R)\alpha = \Delta\mu.\alpha \text{ (say)}.$$

Now let there be two thin prisms, angles α_1 and α_2, and with their refracting edges pointing in opposite directions. The prisms are not necessarily in contact. Then if $\Delta\psi$ is the angle between the red and blue rays after passing through the combination,

$$\Delta\psi = [\Delta\psi]_1 - [\Delta\psi]_2 = [\Delta\mu]_1\alpha_1 - [\Delta\mu]_2\alpha_2,$$

where the suffixes denote the prism for which the quantity in brackets is to be evaluated. If the system is to be achromatic

$$[\Delta\mu]_1\alpha_1 - [\Delta\mu]_2\alpha_2 = 0.$$

This is the condition previously found. But the deviation of the sodium ray is given by

$$\psi = (\mu_1 - 1)\alpha_1 - (\mu_2 - 1)\alpha_2$$
$$= \frac{[\Delta\mu]_1\alpha_1}{\omega_1} - \frac{[\Delta\mu]_2\alpha_2}{\omega_2}$$

where ω_1 and ω_2 are the dispersive powers. If the system is achromatic $[\Delta\mu]_1\alpha_1 = [\Delta\mu]_2\alpha_2$,

so that

$$\alpha_1 = \frac{\psi}{[\Delta\mu]_1\left[\frac{1}{\omega_1} - \frac{1}{\omega_2}\right]},$$

with a similar expression for α_2. Thus it becomes possible to calculate the angles of the prisms so that the system shall be achromatic and the deviation of the mean ray have a given value ψ.

Example.—A combination of prisms of small angle is to be achromatized with reference to the H_α (or C) and H_γ (or F) lines of the hydrogen spectrum. The glasses used have the following refractive indices.

$H_\alpha(\lambda = 6563)$	$D(\lambda = 5893)$	$H_\gamma(\lambda = 4340)$
1·5136	1·5160	1·5262
1·6415	1·6469	1·6724

If the deviation of the mean ray is to be 2°, find the angles of the prisms.

We have $$\omega_1 = \frac{0{\cdot}0126}{0{\cdot}5160} = 0{\cdot}0244\,; \quad \therefore\ \omega_1^{-1} = 40{\cdot}98$$

$$\omega_2 = \frac{0{\cdot}0309}{0{\cdot}6469} = 0{\cdot}0478\,; \quad \therefore\ \omega_2^{-1} = 20{\cdot}92$$

$$\therefore\ \alpha_1 = \frac{2}{0{\cdot}0126 \times 20} = \frac{1}{0{\cdot}126} = 7{\cdot}94^\circ$$

and $$\alpha_2 = \frac{2}{0{\cdot}0309 \times 20} = \frac{1}{0{\cdot}309} = 3{\cdot}24^\circ.$$

[*N.B.*—The equations giving α_1 and α_2, give angles in degrees or radians according as ψ is given in degrees or radians.]

Amici's Prism.—This is a double prism, introduced by Amici in 1860, of the form shown in Fig. 23·14. ADC is a flint glass prism with a right angle at D, and it will be assumed that its refracting angle α_2 is known. ABC is a crown glass prism cemented to ADC along AC by means of a thin layer of Canada balsam. It is assumed that the indices of refraction μ_1 and μ_2 are known for the materials of the prisms ABC and ADC, respectively, for some standard ray—say for sodium light. The problem is to calculate the angle α_1 in order that a ray PQ, incident on AB in a direction parallel to AD, shall emerge parallel to itself, i.e. along a normal to the face CD. Thus the deviation for this ray is zero.

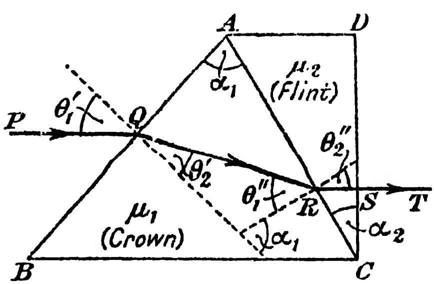


FIG. 23·14.—Amici's Prism.

Let PQRST be the path of such a ray, the angles of incidence and refraction at the different interfaces being as indicated. Then

$$\sin\theta_1' = \mu_1 \sin\theta_2' \quad . \quad . \quad . \quad . \quad . \quad (i)$$

$$\mu_1 \sin\theta_1'' = \mu_2 \sin\theta_2'', \quad . \quad . \quad . \quad . \quad . \quad (ii)$$

and $\theta_2'' = \alpha_2$, on account of the geometry of the system. Also

$$\theta_1' - \theta_2' = \theta_1'' - \theta_2'', \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

since the deviations at the interfaces AB and AC must be equal (and opposite) if ST is to be parallel to PQ. This last equation gives

$$\begin{aligned} \theta_1' &= \theta_1'' + \theta_2' - \alpha_2 \ . \ . \ . \\ &= \alpha_1 - \alpha_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv)} \end{aligned}$$

Thus it only remains to calculate θ_1' in order to determine α_1. Equation (1) gives

$$\sin(\alpha_1 - \alpha_2) = \mu_1 \sin(\alpha_1 - \theta_1'').$$

$$\therefore \ \sin\alpha_1 \cos\alpha_2 - \cos\alpha_1 \sin\alpha_2 = \mu_1[\sin\alpha_1 \cos\theta_1'' - \cos\alpha_1 \sin\theta_1''].$$

$$\therefore \ \sin\alpha_1[\cos\alpha_2 - \mu_1 \cos\theta_1''] = \cos\alpha_1[\sin\alpha_2 - \mu_1 \sin\theta_1''].$$

$$\therefore \ \tan\alpha_1 = \frac{\mu_1 \sin\theta_1'' - \sin\alpha_2}{\mu_1 \cos\theta_1'' - \cos\alpha_2}.$$

Eliminate θ_1'', using $\mu_1 \sin\theta_1'' = \mu_2 \sin\alpha_2$ so that

$$\cos\theta_1'' = +\sqrt{1 - \left(\frac{\mu_2}{\mu_1}\sin\alpha_2\right)^2}$$

$$\therefore \ \tan\alpha_1 = \frac{(\mu_2 - 1)\sin\alpha_2}{+\sqrt{\mu_1^2 - \mu_2^2 \sin^2\alpha_2} - \cos\alpha_2}.$$

Numerical Example.—Take $\alpha_2 = 45°$, $\mu_1 = 1{\cdot}517$ and $\mu_2 = 1{\cdot}650$. There is no need to use the last equation obtained above to find α_1 —we may proceed directly and more simply as follows. We have

$$\begin{aligned} 1{\cdot}517 \sin\theta_1'' &= 1{\cdot}650 \sin 45° \\ \therefore \ \theta_1'' &= 50°\ 17' \\ \therefore \ \tan\alpha_1 &= \frac{1{\cdot}517 \sin 50°\ 17' - \sin 45°}{1{\cdot}517 \cos 50°\ 17' - \cos 45°} = \frac{460}{263} \\ \therefore \ \alpha_1 &= 60°\ 20'. \end{aligned}$$

and θ_1', the angle of incidence on AB is 15° 20′.

The Angular Dispersion for an Amici Prism.—The angle of the two component prisms in an Amici prism are so chosen that a certain ray—generally the sodium ray—travelling in a specified direction shall suffer no resultant deviation after traversing the prism. When a beam of white light parallel to the above specified direction, viz. parallel to the base AD of the second prism, falls on the prism and passes through it, the yellow rays will emerge parallel to AD, but the other rays will be deviated—the white light will have been dispersed. If the emergent beams fall on a converging lens each system of rays will be brought to a focus and a spectrum produced. Thus a spectrum in which the deviation of one particular ray is zero will have been produced. Thus a simple form of direct-vision spectroscope will have been obtained.

The Direct-Vision Spectroscope in Practice.—The single form of this instrument, outlined above, suffers from the defect that the dispersion is small. If, however, two identical Amici prisms are obtained and arranged with the flint prisms adjacent a direct vision spectroscope with greater dispersion is obtained. Such an arrangement of prisms is shown in Fig. 23·15 (*a*). In practice only one flint prism is used in order to avoid the expense of making two separate prisms and cementing them together. It is at once apparent that the mean or standard ray will be undeviated, but although the dispersion is increased it is not doubled as one might expect at first. It may be augmented by increasing the number of prisms, care being taken to arrange them in such a way that the system begins and ends with a prism of crown

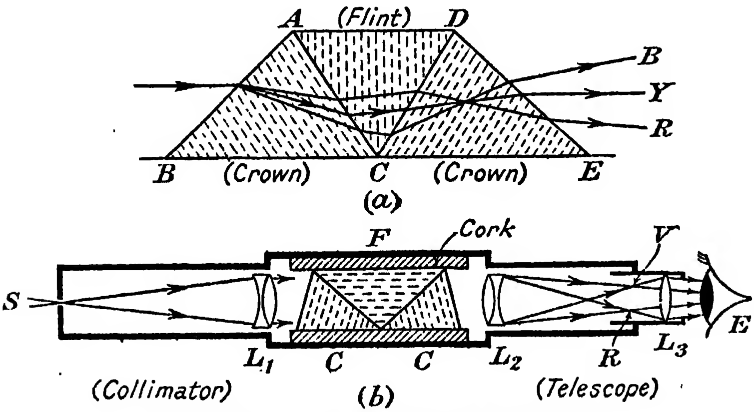


FIG. 23·15.—A Direct-Vision Spectroscope.

glass. The paths of red, yellow, and blue rays are shown in the diagram and it will be noted that the rays cross one another in the second crown glass prism. This fact can only be verified in a given instance by numerical calculation.

The main optical features of a direct-vision spectroscope are shown in Fig. 23·15 (*b*). The three prisms are mounted in a brass tube and held in position with the aid of pieces of cork. S is a narrow slit parallel to the refracting edges of the component prisms ; it lies in the first focal plane of the achromatic converging lens L_1. Thus a parallel beam of white light falls on the system. The emerging systems of parallel beams fall on another achromatic converging lens L_2 and are brought to a focus in its second focal plane to form a spectrum VR. This is examined with a lens L_3 which acts as a simple magnifying-glass. Thus the spectrum may be observed. The collimator and telescope must be focused in the

usual way before the instrument can be used. The instrument is a convenient one for examining flame spectra, but its resolving power is small, i.e. lines very close together cannot be seen separated.

Colour due to Absorption.—Absorption plays a very important part in determining the colour of natural bodies. We have already seen that when white light passes through a body the emergent light is coloured. This is due to the fact that in its passage through the body some of the constituents of the white light have been absorbed, so that the colour of a body as observed by transmitted light is really a composite effect due to all those colours not absorbed by the body.

The colour of an object seen by reflected light is determined by the nature of the light scattered from it and received by the eye. It is believed that the molecules are the entities responsible for this scattered light. Now it is not only the molecules at the surface which take part in this scattering action but also some which are inside the body. If absorption occurs within the body the light scattered by these will be robbed of some of its constituents so that the tint of the body is again a resultant effect. It is interesting to note that many substances which are coloured when they exist in large pieces appear white after they have been crushed to a very fine powder. This is because even when the substance has been reduced to powder form the crystalline character of the substance is still retained so that millions of tiny facets are present which reflect the light and do not permit an appreciable amount of light to penetrate into the interior. Thus crushed ice appears white, although ice in bulk is transparent owing to its exceedingly small reflecting power. Similarly red and blue glasses, crystals of copper sulphate, etc., all tend to become white when powdered. Again, beer and other beverages have a definite colour when seen in bulk and yet the froth on them is white. This is because the thin liquid film forming each bubble of the froth reflects most of the incident light and so it appears white. If such a froth is examined in a red or a yellow light it assumes the colour of the light by which it is examined.

The examination of the colours of bodies when they are illuminated by monochromatic light is very instructive. Thus a piece of white paper appears red when placed in a red light, green in a green light, etc. This is because the paper reflects light of all colours. On the other hand, a red flower only appears red when viewed in white or in red light. If such a flower is examined by blue light it appears black since it can only reflect red light.

When a sodium flame, a source of monochromatic light (yellow), is viewed through an ammoniacal solution of cuprous chloride placed between the eye of an observer and the sodium flame this

latter cannot be seen, for the yellow light of the flame is strongly absorbed by the blue solution, although white objects having a blue tint may be discerned. We might therefore expect that when yellow and blue pigments are mixed the colour of the mixture would be black: actually it is green. This is because the colours of the pigments are not pure but only appear to have a definite colour corresponding to that which they most copiously reflect. Thus the blue pigment absorbs red, orange, and the yellow rays, whilst the yellow absorbs the blue and violet. The only colour not absorbed by either pigment is green, so that this is the colour of the mixture when examined in white light.

The Mercury Vapour Lamp.—Another convenient source of monochromatic light for use in experiments similar to the above is the mercury vapour lamp. A simple form due to WARAN is shown in Fig. 23·16. The apparatus is made of pyrex glass and has two tungsten electrodes sealed in at E and F. These are surrounded by crucibles containing mercury so that they may be connected through an adjustable resistance [about 80 ohms] to a 110-volt D.C. supply. The portions A, B, and part of C are filled with mercury and the lamp connected at H to a water pump. When the pressure in the apparatus is sufficiently low the mercury in B begins to descend and finally breaks at K. The energy of the spark produced at K when the current is thus broken volatilizes some of the mercury and the vapour formed carries the current, the lamp emitting a strong green light. The light is not monochromatic for, in addition to the green, it contains yellow, blue and violet rays. These may be removed by passing the light through an aqueous solution of malachite green containing potassium bichromate. The amounts of these substances must be adjusted by trial. [For details concerning A, cf. p. 98.]

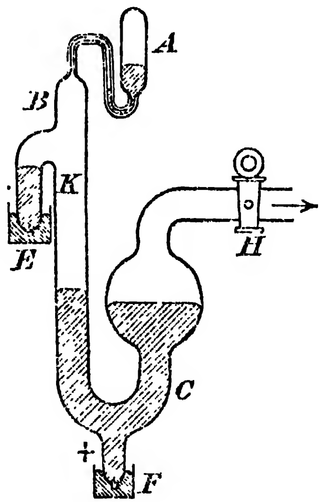


FIG. 23·16.—Mercury Vapour Lamp.

The lamp remains in action almost indefinitely, especially if E is the cathode: this is because positive mercury ions leave the anode and, having lost their charge, collect above E. If E is the anode the mercury in the limb E disappears and the lamp may cease to work.

In addition to the above-mentioned visible rays, the lamp is

also a source of ultra-violet rays. These rays are freely transmitted if the walls of the vessel are of silica. Such lamps must not be observed directly on account of the painful sensation at the back of the eyes experienced some time after the retina has had ultra-violet light incident upon it. Smoked glasses should be worn, for safety, with all experimental mercury lamps—the glass alone is sufficient to absorb the harmful rays, but the glass is made dark to reduce the intensity of the visible rays.

Modern industrial forms of mercury vapour lamps are shown in Fig. 23·17 (*a*) and (*b*). In each the cathode is a pool of mercury, C, and the anode an iron ring, A. When the tube is straight and in a horizontal position, the mercury connects A and C. When the tube is connected to d.c. mains through an adjustable resistance, R, a current flows through the mercury when the switch is closed. On raising the end A of the tube, the mercury is broken at a point in the tube. The small arc formed at the point of rupture volatilizes some of the mercury and a discharge is maintained in the tube.

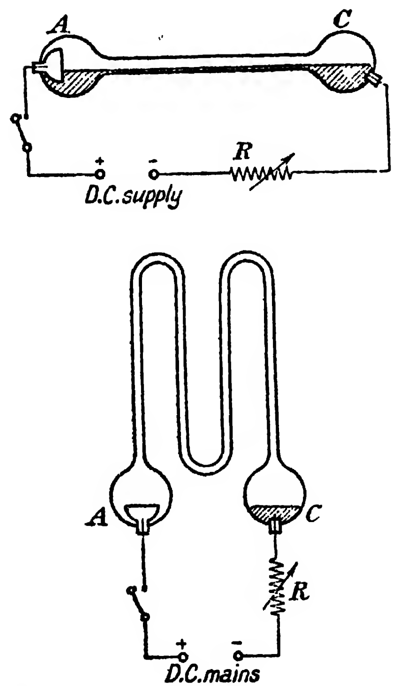


FIG. 23·17.—Mercury Vapour Lamps—Modern Industrial Forms.

In type (*b*), the initial discharge is obtained by applying a high potential difference to the tube; this is then automatically cut out and the d.c. supply connected before the discharge has ceased. The lamp continues to work.

These lamps may also be designed to work from a.c. mains—electrical devices are used to maintain the arc during a certain portion of each cycle when the p.d. is zero or insufficient to operate the tube.

If the light from a mercury vapour lamp falls on a person having a ruddy countenance a very ghastly effect is produced, for the red portions appear black while the white portions appear green.

Radiations beyond the Visible Region of the Spectrum.—According to the wave theory of light all light sensations are a manifestation of different vibrations in the æther. Each different

vibration is characterized by a definite wave-length and by methods we shall describe later it is known that the red waves are longer than the violet ones, and that the intermediate colours in the spectrum are associated with wave-lengths intermediate between those of red and blue. We have already seen [cf. p. 327] that beyond the red end of the spectrum there are radiations characterized by their heating effects. These are the so-called *infra-red* rays. A simple method for showing that the heating effect of the radiations from an electric lamp is not confined entirely to its visible rays is as follows. The radiation is allowed to fall on a thermopile connected to a sensitive galvanometer: the deflexion is noted. When a piece of sheet glass is placed between the source and thermopile the deflexion is reduced considerably. Now the amount of visible light passing through the glass scarcely differs from that which is incident upon it: the glass must absorb therefore some radiations which were responsible for heating the surface of the thermopile, i.e. the source does emit infra-red rays.

In 1800, Sir W. Herschel established the existence of rays in the solar spectrum which are definitely 'less refrangible than any of those that affect the sight.' They could heat, but not illuminate bodies, and this explained why they had previously escaped notice. In these experiments he used the blackened bulb of a thermometer to detect the heat rays: absorption of these rays caused the temperature of the bulb to rise. Finally he investigated the heating effect in the different parts of the visible spectrum and found that it was a minimum at the violet end, the intensity gradually increasing towards the red and reaching a maximum in the region beyond the red.

Leslie challenged the accuracy of Herschel's experiments and denied completely the existence of these invisible heat rays. He said that if Herschel's results were correct, then, as a consequence, the heat focus of a burning lens would be different from its light focus, which, Leslie maintained, was contrary to experience. We now know that Leslie was wrong.

In 1802 Young said, 'at present it seems highly probable that light differs from heat only in the frequency of its undulations or vibrations; those which come within certain limits, with respect to frequency, being capable of affecting the optic nerve, constitute light; and those which are lower, and probably stronger, constitute heat only.' Later he described Herschel's discovery of the less refrangible heat rays as one of the greatest since the time of Newton.

Bérard confirmed, in general, the work of Herschel, but his experiments showed the maximum heating effect to be in the

region of the extreme red rays, whereas Herschel's maximum was in the infra red region.

Seebeck showed that the position of the maximum was affected by the nature of the material of the prism. Some of his results are shown in the following table:—

Position of Maximum.	Kind of Prism.
Yellow	Water
Orange	Clear solution of sal ammoniac or of corrosive sublimate sulphuric acid (concentrated).
Red	Crown glass
On the limit of the red . .	Prism containing lead and coloured yellow
Beyond the red	Flint glass

These experiments explained why the earlier workers had obtained conflicting results.

In 1840 Sir J. Herschel announced his discovery of a method whereby the heat rays of the solar spectrum could be made to leave a visible trace. The solar spectrum was projected on an extremely thin piece of paper, blackened evenly with soot on the back surface and moistened with alcohol on the front or exposed surface. Unequal heating produced unequal evaporation and where the paper dried it became light in colour. Evaporation only occurred over certain regions—at others, the paper was still moist; hence this portion of the spectrum was not continuous but contained absorption bands. With better apparatus these bands have been shown to consist of many lines, i.e. Fraunhofer lines have been shown to exist in the infra-red region of the solar spectrum.

Since 1882 this region has been very carefully explored and the wave-lengths of the lines in it measured. It is important to note that glass is not transparent to the heat rays of longer wavelength so that prisms of rock salt have to be used. Even these absorb the still longer rays so that diffraction gratings [cf. p. 556] ruled on speculum metal have to be employed. Modern work has shown that a study of absorption bands in the extreme infra-red region is most important, for it gives us a clue to the structure of the molecule of the absorbing material.

A typical apparatus for exploring the infra-red region is shown in Fig. 23·18. Light from a source, O, is focused on a slit, S_1, by means of a concave mirror, M_1. S_1 is in the focal plane of a concave mirror, M_2, so that the light is reflected from this mirror

as a parallel beam, which falls on a prism, ABC, made of rock salt—or a grating may be used if rock salt absorbs heat rays. The refracted beams fall on a concave mirror, M_3, in the focal plane of which lies the blackened surface of a thermopile, T, connected to a sensitive galvanometer, G. When the mirror M_3 is rotated about a vertical axis, various 'lines' in the spectrum fall on T and the galvanometer is deflected. The greater the deflexion, the greater the intensity of the line.

[In work of this kind, the thermal capacity of the thermopile must be small so that the rise in temperature of the junctions shall be large; moreover, the actual junctions must lie on a straight line so that only one line in the spectrum shall fall on them at a time—cf. p. 324.]

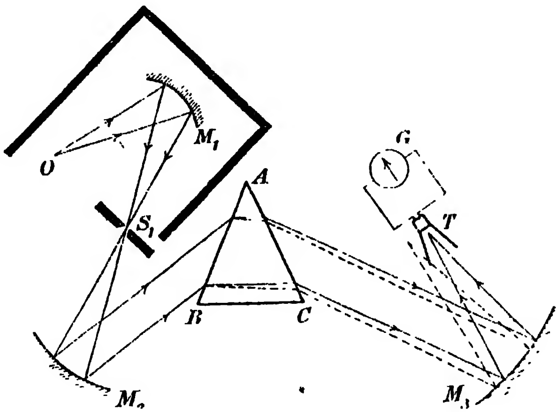


FIG. 23·18,—Apparatus for Exploring the Infra-red Region of the Spectrum.

Beyond the violet end of the spectrum there are the ***ultra-violet*** rays which may be detected by allowing the spectrum of the light from an arc lamp to fall on a barium platino-cyanide screen, which then glows with a green light where the ultra-violet rays strike it. Since glass absorbs the shorter ultra-violet rays the effect is increased by using a quartz prism. These rays have wave-lengths decreasing in magnitude as the rays become further removed from the visible spectrum. Most frequently these rays are detected photographically, for they possess the property of darkening a photographic plate when one is exposed to the rays and then developed. Ultra-violet rays may also be detected by the photo-electric effect they produce [cf. p. 477].

The existence of ultra-violet light in the spectrum of the light from an arc lamp, for example, may be demonstrated very strik-

ingly as follows. Suppose that RV, Fig. 23·19, is the visible spectrum produced in the usual way. Let Z be a screen having a slit at W. Let $A_1B_1C_1$ be another prism placed as shown, and P a barium platino-cyanide screen. A patch of light appears on P

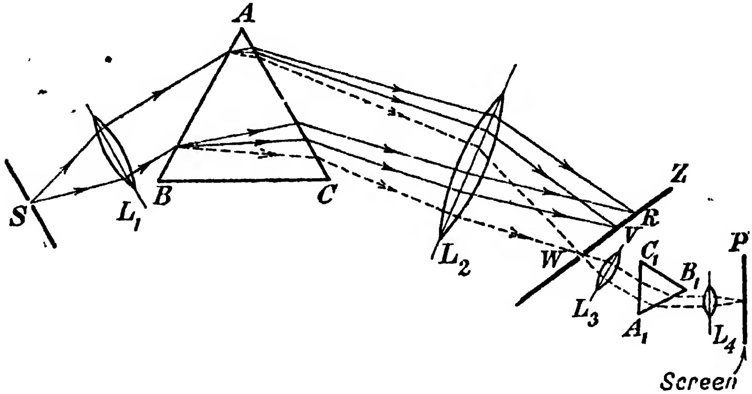


FIG. 23·19.—Detection of Ultra-Violet Light beyond the Violet end of the Visible Spectrum.

when all the component parts of the apparatus have been properly adjusted. This shows the existence of radiations beyond V and that they are refracted further when incident upon another prism. If quartz prisms and lenses are used the existence of radiations still further out from V may be shown.

A Comparison of the Spectra of a Red-hot Iron Ball and of an Iron Arc.—The spectrum of the heated ball may be obtained by placing it in the position O, Fig. 23·4, p. 450. It will be found that the red end of the spectrum is the most intense. To detect any radiations in the infra-red region, the spectrum is allowed to fall on the surface of a thermopile connected to a sensitive galvanometer. It is advisable to place a slit immediately in front of the thermopile so that only a narrow region of the spectrum is investigated at once. As different parts of the spectrum are examined in turn it will be noticed that the deflexion of the galvanometer first becomes appreciable in the yellow region, increases to a maximum as one passes through the red to the infra-red region, and then decreases. On replacing the thermopile by a barium platino-cyanide screen no ultra-violet rays will be detected.

If the above experiment is repeated using an iron arc (i.e. iron instead of carbon electrodes) the yellow and green regions are most intense when examined visually. A thermopile shows that the intensity of the infra-red rays has increased and that the maximum heating effect is nearer to the blue end of the spectrum. At the other

end a fluorescent screen shows that the spectrum is very rich in ultra-violet light—in fact, it is advisable never to look directly at this arc, for the action of these rays is most harmful, and although no immediate effect is noticed, a person who has been affected in this way often experiences a terrible pain at the back of the eyes about 2 a.m., i.e. when the vitality of the body is lowest. Although ultra-violet rays always produce this effect, many now believe that it is due in part to the infra-red rays as well.

An important difference between the two spectra now under discussion is that the spectrum formed by the light from the red-hot ball is continuous, whereas the light from the iron arc gives rise to a spectrum consisting of many bright lines on a faint continuous background.

Experiment.—Project an arc-light spectrum in a darkened room on a sheet of photographic printing-out paper [P.O.P.]. After a short time the paper becomes darkened, but not uniformly. The maximum blackening occurs in the extreme violet or ultra-violet, while the red and yellow rays produce practically no effect. Hence if the light is passed through a piece of red glass before it reaches the P.O.P., no darkening occurs. Thus photographic plates and papers may be manipulated with safety in a red light, unless the plates happen to be panchromatic ones [cf. p. 533], when complete darkness is necessary.

Tyndall's Experiment.—Since infra-red rays are refracted when they pass from one medium to another [except at normal incidence], they should be capable of being focused by a lens. Sunlight was passed through a solution of iodine in carbon bisulphide, which is opaque to visible radiations but transmits infra-red rays freely. The emergent light was focused by means of a rock-salt lens on a piece of blackened platinum foil. This soon became red hot.

The Photoelectric Effect.—Hallwachs discovered that when a clean piece of zinc was insulated and charged negatively, it lost its charge when illuminated with ultra-violet light, but that the charge was retained if it were positive. It is now assumed that the particles of negative electricity—the electrons—are released by the ultra-violet light and that these are repelled from the negatively charged zinc.

Phosphorescence and Fluorescence.—Whenever radiation falls on a body and the sum of the transmitted and reflected energy is not equal to the energy in the incident radiation, it follows that the difference must have been converted into some other form of energy; generally the body is heated. Sometimes, however, the rise in temperature is small and yet the body emits visible radiation differing in wave-length from that of the incident light. This peculiar emission may cease immediately the exciting agent is

removed—the body is said to fluoresce and the phenomenon is spoken of as that of ***fluorescence***. If the emission continues for some time, however, we have the phenomenon of ***phosphorescence***. Probably there is no sharp line of demarcation between these two phenomena.

Calcium sulphide [Balmain's luminous paint] and native zinc sulphide, after exposure to sunlight or the light from an arc lamp, are found to glow with a bluish or greenish light respectively when removed to a darkened room, i.e. they phosphoresce. If, however, the paint is gently warmed over a bunsen flame the rate of emission of the phosphorescent light is greatly increased for a short time and then ceases. When the paint is subjected to sunlight once more, phosphorescent light is again emitted. Similarly calcite [native $CaCO_3$] phosphoresces with a strong red light when exposed to an intense beam of cathode rays.

As an example of fluorescence we may cite that of an aqueous solution of quinine sulphate, which BREWSTER found to emit a vivid blue light in all directions when exposed to sunlight. Or again, if, in a darkened room, a beam of light is passed through water containing a few drops of an alcoholic solution of fuschine, a green fluorescent light indicates the path of the beam.

Stokes' Law.—Let us suppose that the phosphorescence of some Balmain's paint has been destroyed by a gentle application of heat and that the paint is then spread in a darkened room on a strip of paper. To excite the paint again, let the spectrum of an arc lamp be cast upon it for several minutes. On removing the arc lamp it will be found that the maximum phosphorescence occurs where the paint has received the violet and ultra-violet rays, and that the red rays have produced no effect. This experiment proves that it is the most refrangible rays, i.e., the rays of short wave-lengths which are responsible for the excitation of phosphorescence. In addition, the colour of the emitted light is bluish green, whereas the exciting rays are violet and ultra-violet. This is merely a particular instance of a general law discovered by STOKES. It states that the wave-length of the fluorescent or phosphorescent radiation is always greater than that of the exciting light. This change in colour may be demonstrated as follows :—The rays from an arc lamp are concentrated on a piece of canary glass [one which contains uranium oxide] and the orange-yellow and green rays intercepted by a piece of dense cobalt glass. The canary glass shines with a vivid green light although the exciting rays are blue and violet.

The above experiment becomes more striking if the blue glass is replaced by a piece of 'Ultra-Violet' glass [manufactured by Messrs. Chance, of Birmingham]. This is a special glass absorb-

ing practically the whole of the visible rays but being very transparent to ultra-violet radiations. These excite the canary glass to fluorescence.

Similarly, if egg-shells cooled to a temperature below − 100° C. by means of a mixture of solid carbon dioxide and ether, or by liquid air, are placed in a beam of ultra-violet light, they fluoresce.

When a few drops of an alcoholic solution of eosin are added to water placed in the path of the light from an arc lamp the fluorescent light indicates the path of the beam. As more of the solution is added the fluorescence becomes most vivid on the side where the rays enter the solution. This is because the exciting radiation [the violet rays] is gradually absorbed within a short distance and none remains to produce further fluorescence. A spectroscopic examination of the transmitted light shows that the violet light originally present in the light from the arc is missing.

In all instances the fluorescent light is less refrangible than the exciting light. In recent years apparent exceptions to this rule have been reported, but they have been proved to be not genuine.

Methods of Detecting Fluorescence and Phosphorescence.—(*a*) ***Stokes' method for detecting fluorescence.*** When the fluorescence is feeble it may be masked by the effect of the exciting beam. To avoid this, Stokes made use of the facts stated in the law which bears his name. The substance under

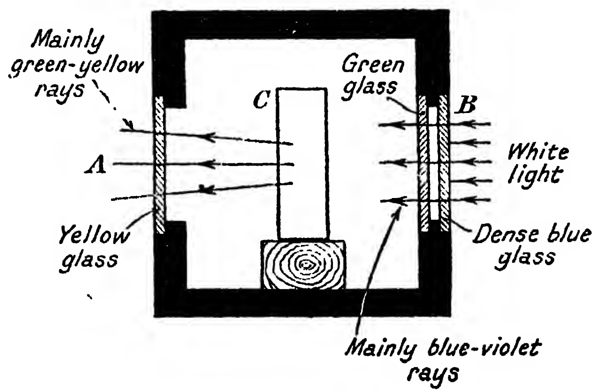


Fig. 23·20.—Stokes' Method for Detecting Fluorescence.

examination was placed at C, Fig. 23·20, in a box whose interior had been blackened. A and B are two apertures in the box. B was covered with two sheets of glass, one green and the other dark blue so that if a source of white light was placed just outside B all the orange-yellow and red rays were intercepted. On the other hand, A was covered with a piece of yellow glass which completely

absorbed any blue-violet rays. Hence, on looking through A, the field of view would be dark and would remain thus when a substance which did not fluoresce was introduced into the box at C. If, however, fluorescence was excited, the body at C emitted green light and since this was transmitted by the yellow screen, the object was rendered visible.

(*b*) ***Becquerel's Phosphoroscope.*** To overcome the difficulty that the light from a fluorescent body vanishes in a small fraction of a second after the exciting light has been removed, BECQUEREL devised the phosphoroscope shown diagrammatically in Fig. 23·21 (*a*). It is only suitable for use with transparent

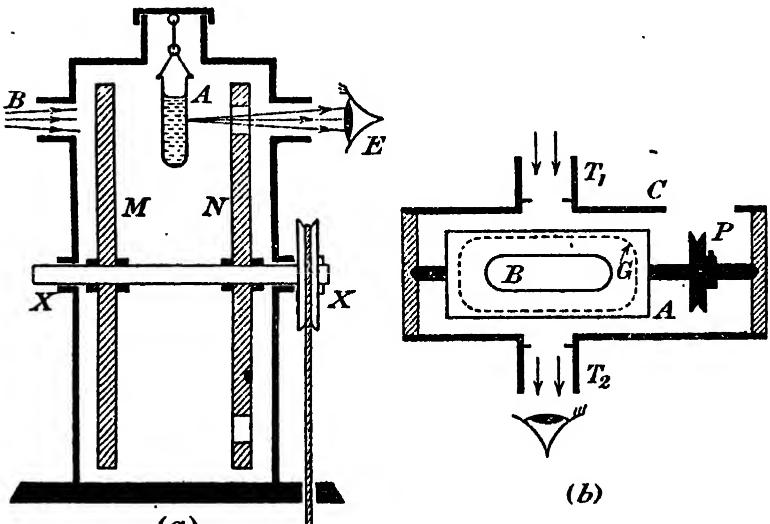


FIG. 23·21.—(*a*) Becquerels' Phosphoroscope (for Transparent Substances). (*b*) Lenard's Cylindrical Phosphoroscope (for Solids and Powders).

substances. The substance to be tested, if it be a liquid, is placed in a tube at A and viewed by an observer at E, the exciting radiation entering the chamber through an aperture at B. M and N are two circular metal discs capable of rotating rapidly round a common axis XX. Each disc is pierced by an equal number of regularly spaced peripheral holes, which are so disposed that a hole in M passes A completely before the corresponding hole in N comes directly between A and E. Thus A is illuminated by an intermittent beam of light and since when it is viewed at E it is receiving no light from the source it will only be seen if it phosphoresces. Since the angular distance between a hole in M and a corresponding one in N and the number of revolutions per second of the discs

are known, the duration of the phosphorescent light is determinable. Becquerel discovered substances which phosphoresced for a few thousandths of a second, and that the light from fluorescent liquids remained for such a short period after the exciting agent had been removed that he could not measure the duration, however quickly the discs revolved.

(c) ***Lenard's Cylindrical Phosphoroscope.***—This instrument was invented in 1888 especially for the examination of the time of duration of the phosphorescent light from solids. A, Fig. 23·21 (b), is a brass cylinder capable of rotation about a horizontal axis—the pulley P enables this to be done. G is a glass or quartz container inside A: it contains the solid and is held in position by means of light springs not shown. The whole is supported inside a metal tube C provided with side tubes T_1 and T_2. The exciting radiation enters through T_1 and any phosphorescent effect is observed by an observer at E. The cylinder A has an elliptical aperture B, so that the substance is illuminated when B is opposite T_1: when B is opposite T_2 any phosphorescent effect can be observed. Its duration is one-half the period of rotation of the cylinder. This apparatus is suitable for use at different temperatures.

Rainbows.—These are produced when sunlight falls on rain or the spray from a waterfall, although to be seen the observer must stand with his back to the sun. As a rule, two rainbows may be

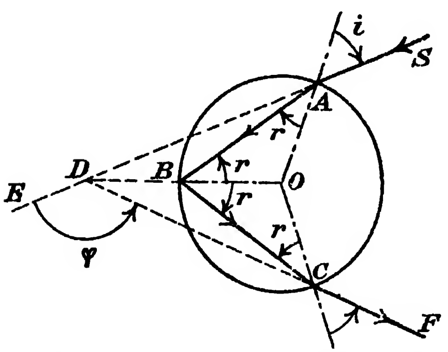


Fig. 23·22.—Path of Light Ray suffering one Internal Reflexion in a Rain Drop.

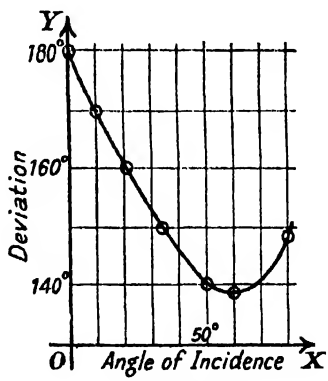


Fig. 23·23.

seen: they are known as the ***primary*** and ***secondary*** rainbow respectively. The primary is the brighter, and is produced by rays of light which have suffered one internal reflexion in the raindrops. The secondary rainbow is formed by rays which have undergone

two internal reflexions. To account for the formation of the primary bow let us calculate the deviation, ϕ, when a ray SA, Fig. 23·22, is incident on a sphere the material of which has a refractive index μ, and this ray finally emerges along CF after one internal reflexion at B. If O is the centre of the sphere

$$\phi = \pi - 2\widehat{ADO} = \pi - 2[\pi - (i - r) - (\pi - r)] = \pi + 2i - 4r.$$

Fig. 23·22 has been constructed on the assumption that the light is monochromatic. If it is not, dispersion will take place when the ray enters the drop so that coloured rays emerge.

Fig. 23·23 shows the relation between the angle of incidence and the angle of deviation for red rays. It shows that for an angle of incidence of about 60° the deviation is a minimum and it is only when the rays traverse the drop in such a way that the deviation is a minimum that they are sufficiently concentrated in one direction to be seen.

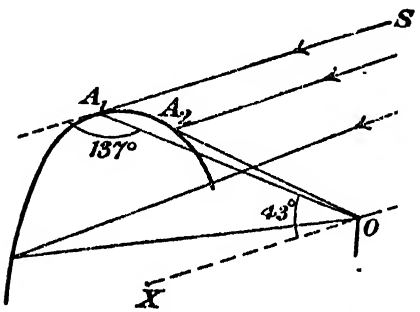


FIG. 23·24.—The Rainbow.

Fig. 23·24 shows an observer OP with his back to the sun facing the raindrops A_1, A_2, etc. If through O a cone having a semi-vertical angle of 43° and its axis OX parallel to the sun's rays is considered, all the red rays emerging from drops lying on the surface of this cone will travel towards O along the generators of this cone. Since all such rays will have suffered minimum deviation they will easily be seen. Now the refractive index of water for violet light is greater than that for red, so that the angle of minimum deviation is greater for violet rays than for red. Hence the semi-vertical angle of the corresponding cone will be less than for the red rays. The primary bow therefore consists of a prismatic arc coloured red on its outside.

A similar argument shows that the secondary bow is coloured violet on its outer margin.

EXAMPLES XXIII

1.—Explain how a narrow slit, a prism, and two convex lenses, may be arranged to produce a pure spectrum. Show also, how it is possible to arrange two or more prisms to produce (*a*) dispersion without deviation of the mean ray, (*b*) deviation without dispersion.

2.—Describe the optical system of a spectrometer and state how you would use it to find the refractive index of water. Prove the more important formulæ you would use.

3.—Define *chromatic dispersion* and explain how it is possible to obtain achromatic prisms and lenses.

4.—Write an essay on colour.

5.—Describe experiments to show that the spectrum of an iron arc extends beyond the limits of the visible region. Discuss how the results would be modified if a red-hot iron ball were used as a source of radiation.

6.—Two glasses have dispersive powers in the ratio 3 : 2. These glasses are to be used in the manufacture of an achromatic objective of focal length 20 cm. What are the focal lengths of the lenses ?

7.—An achromatic telescope objective of 100 cm. focal length is to be formed with two lenses made of the glasses specified below. Find the focal length of each of these lenses, stating whether it is divergent or convergent.

	μ red.	μ blue.
Glass A	1·5155	1·5245
Glass B	1·641	1·659

(N.H.S.C. '29.)

8.—Establish the necessary and sufficient condition that two thin lenses in contact may form an achromatic combination.

A converging achromatic lens of 50 cm. focal length is to be constructed from the following glasses :—

Type of glass	μ_D	$\mu_F - \mu_C$
Crown	1·520	0·0087
Flint	1·615	0·0166

If the converging lens of the combination is to have surfaces of equal curvature and to be cemented to the other component lens, find the radii of curvature of the faces of the lenses.

9.—A converging achromatic lens of focal length 100 cm. is to be constructed from the following glasses :—

Type of glass	μ_D	$\mu_F - \mu_C$
Crown	1·516	0·0081
Heavy flint	1·647	0·0192

(*a*) If the diverging lens has one face plane, calculate values for the radii of curvature of the other faces of the lenses.

(*b*) If the converging lens of the combination is to have surfaces of equal curvature and to be cemented to the other component lens, find the radii of curvature of the faces of the lenses.

CHAPTER XXIV

PHYSIOLOGICAL OPTICS

The General Structure of the Human Eye.—Vision is the sense by which we are enabled to form a mental picture of external objects. It is now believed that light consists of waves, which are the stimulus whereby the retina is excited. The sensations produced upon the retina enable persons and animals to judge colour, estimate distances and, in general, to form some idea of the external world.

In Fig. 24·1 there is reproduced a section of the human eye. Considered in a very general manner the eye consists of an almost

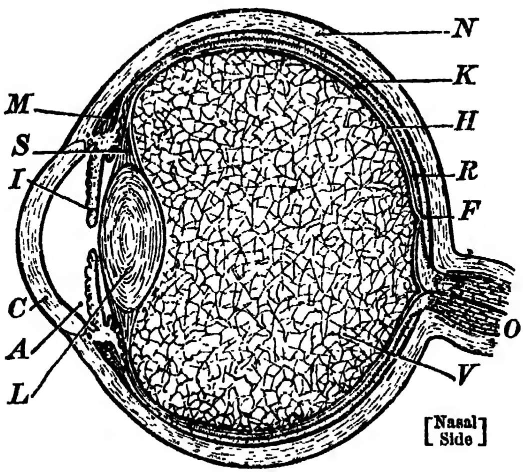


FIG. 24·1.—Horizontal Section through the Human (Right) Eye.

spherical chamber which is provided with an aperture through which the light vibrations pass. The whole is known as the ***eyeball*** and is contained in a cavity of the skull called the ***orbit.*** In front of the eyeball are the ***lids*** and lacrimal apparatus. The transparent anterior part, C, of the eyeball is called the ***cornea,*** whereas the posterior portion, which is opaque and envelops about five-sixths of the eyeball, is called the ***sclerotic,*** N. The cornea, C, is really a protuberance on the eye, so that its radius of curvature is less than that of any other part of the eye chamber. Immediately in front of the sclerotic, or outer covering of the eye, is the ***choroid,*** K;

this tissue has a liberal supply of black pigment cells on its internal surface, and these cells absorb any superfluous light which may enter the eye. The interior surface of the posterior portion of the eye is the ***retina*** or ***arachnoid***, R [so named on account of its structural resemblance to a spider's web].

Near to the point where the sclerotic merges into the cornea is situated the contractile membranous diaphragm called the ***iris***, I. This is generally coloured, a fact which is used by credulous people in their efforts to gain an insight into the future.

The aperture in the iris is called the ***pupil***, and this is not situated at the geometrical centre of the iris, but is slightly displaced towards the nasal side of the eye.

Immediately behind the iris is the ***crystalline lens***, L, which is supported from the walls of the eye by means of an annular diaphragm, called the ***suspensory ligament***, S. It is formed of a non-contracting tissue. In close proximity to this ligament is the ***ciliary muscle***, M, which is, of course, shaped like a ring; its action is governed by the ciliary nerves. Physiologists have been able to show by experiment that this muscle actually pulls at the point where the sclerotic is attached to the cornea.

The portion of the eyeball between the lens and retina is filled with a transparent gelatinous substance, known as the ***vitreous humour***, V. It consists very largely of water, containing traces of proteids, organic and inorganic salts. The hyaloid membrane H encloses the vitreous humour. In front of the lens is the ***aqueous humour***, A, a watery liquid containing a minute trace of sodium chloride.

At the back of the eye, with its centre at the point where the line of sight intersects the retina, is the ***yellow spot*** or ***macula lutea***, F. The depression at the centre of this spot is termed the ***fovea centralis***, and it is here that the acuity of vision is greatest.

The Cornea.—The protuberance on the eyeball which is exposed to view when the eyelids are opened is the cornea, and it was originally believed to be spherical. It is now known that its shape is much more complicated; it is more flattened above than below, and more flattened on the nasal side than on the temporal side. Such facts have been ascertained by means of an instrument which is called the ophthalmometer.

The Crystalline Lens.—The eye lens is not symmetrical about a plane which passes through its periphery; the part which faces the retina has a smaller radius of curvature than the anterior portion. In addition the lens is not homogeneous; it consists of many layers which become more dense, i.e. the refractive index increases, as they approach the inner regions of the lens. Such an arrangement as this tends to improve the sharpness of the image.

The Retina.—This transparent membrane, which is excited by the light stimulus, forms five-sixths of the posterior inner surface of the eye. The part of the retina which is in contact with the vitreous humour consists of a thin layer of connective tissue; the more remote side of the retina is also composed of a layer of this tissue and the two layers are joined together by more tissue. The ***optic nerve,*** O, which conveys the messages excited by the light impacts upon the retina to the brain, passes through the retina at a point situated on the nasal side of the eye. From this point the nerve spreads itself out along the retina, and some nerves fill the intervening spaces in the central portion of the retina. These features are characteristic of the complete retina with the exception of the part which is pierced by the optic nerve—this is called the ***blind-spot,*** about which more must be said subsequently. The inner layer of the retina is rich in nerves and, at certain points, these emanate from large cells called the ***ganglionic corpuscles*** [cf. Fig. 24·2]. Attached to the transverse bundles of connective tissue, which cross the retina, comes the remarkable layer of rods and cones —the ***bacillary layer*** or ***Jacob's membrane.*** This membrane, in which the rods are more numerous than the cones, is followed by a layer of pigment, beyond which the limiting layer of the choroid is encountered.

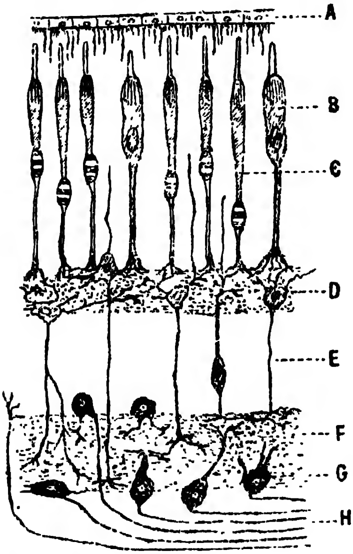


FIG. 24·2.—Diagram of the cell layers in the retina (highly magnified).

A. Pigment cells. B. Cones. C. Rods. D, F. Felt-work of dendrons. E. Axon of one of the cells which lie between the rods and cones and G, the ganglion cells. H. Axons passing from ganglion cells to optic nerve. (After Stöhr.)

The rods have a twofold structure—the inner and outer limbs respectively. The inner limb is easily stained by reagents such as carmine; the outer limb is formed of a highly refracting medium which is not readily stained. It has a pinkish colour and is very sensitive to light radiation; its volume increases under the influence of light but it resumes its original volume when the source of radiation is removed. The pink colouring matter is known as the ***visual purple*** or ***erythropsin*** because the pigment is soluble in certain reagents producing a purple solution, which is readily bleached in daylight.

The cones, also, have a double formation; the inner limb

resembles the inner limb of a rod, but the outer limb is conical in shape, and contains no visual purple.

The Visual Purple.—The visual purple, or erythropsin, is destroyed by acids, alcohol, chloroform or caustic soda. It has already been stated that the visual purple is bleached by light, but this action can be retarded by the addition of a 4 per cent. solution of alum. Using this fact KÜHNE succeeded in obtaining a photographic image upon the retina of a rabbit's eye. The pupil of the eye having been enlarged by dosing the animal with atropine, it was placed in front of a window for a few minutes and then destroyed. The retina was then obtained and washed in the above alum solution. A clear image of the window was visible even after the lapse of several days.

When the visual purple was discovered [and it is found in the rods of the eyes of many animals] it was thought that it was the ultimate means of detecting light. It is now known, however, that snakes and some birds only have cones, so that the ultimate organ of sight is still a mystery.

The Blind Spot.—Owing to the fact that the retina has been pierced by the optic nerve at one spot it is not surprising to find that this region is insensitive to a light stimulus. It is known as the ***blind spot*** and, like the yellow spot [cf. p. 485], is about 0·25 cm.

FIG. 24·3.

in diameter. Close the left eye, gaze intently at the small cross in Fig. 24·3 and, commencing with the book about 40 cm. away, gradually move it towards the eye. Suddenly the small black disc cannot be seen—its image falls on the blind spot, so that its existence is not discerned. The positions of the spot and cross must be reversed if the right eye is closed.

The Formation of Retinal Images.—If an object AB, Fig. 24·4, is placed at a distance in front of the eye, a real inverted image A′B′ is formed upon the retina.

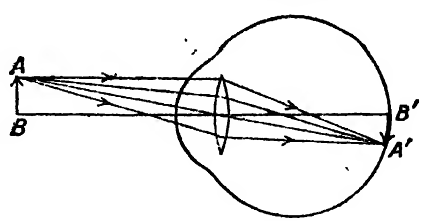


FIG. 24·4.—Formation of an Image by the Eye.

Experiment.* — *To prove that images on the retina are inverted.—A pin-hole is made in a piece of postcard and held 3 cm. in front of the eye and towards a white background. Since the first focal point of the eye is 3 cm. in front of it, the rays entering the eye are parallel. If an object is placed between the hole and the eye an erect

shadow is thrown on the retina. But when a pin, with its head uppermost, is held in this position, it appears to be inverted. Also, when it is moved across the field—it appears to move in the reverse direction. It is therefore clear that in interpreting our sensations an inverted image on the retina is regarded as if it were upright.

The Excitation of the Retina.—The interpretation of a light stimulus by the retina depends to a very large extent upon the intensity of the light as well as upon its duration. A lightning flash is easily seen although its duration is small; on the other hand, the dark green lamp which is used in the development of panchromatic plates cannot be seen on first entering the dark room, yet it and the objects in its immediate neighbourhood become very distinct after a few minutes, and the light appears so intense that one wonders that the plates are not fogged.

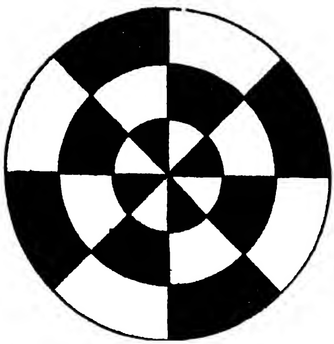


FIG. 24·5.—Helmholtz's Wheel.

An experiment, which follows, also proves that the sensation is not immediately interpreted by the brain, and that the interpretation persists after the removal of the exciting stimulus. Helmholtz's wheel, Fig. 24·5, is made to rotate about its centre. When the wheel is rotating very slowly, the black and white sectors are distinct; when the speed increases, the transverse edges tend to become blurred. This phenomenon proves that the interpretation of the sensation is delayed. A still more rapid rotation of the disc and the interpretation has not sufficient time to wane to zero before it is stimulated again—the disc becomes grey all over but the colour is not uniformly grey. When the wheel rotates yet more rapidly the light stimuli follow so swiftly that the disc appears uniformly grey.

Accommodation.—The great difference between the eye, as an optical system, and a biconvex lens lies in the fact that in the eye the distance between the lens and retina [the seat of the image] is invariable—in an optical system such as that found in a precision camera this is not so. An eye which, when at rest, i.e. without strain, can clearly discern a remote object, is

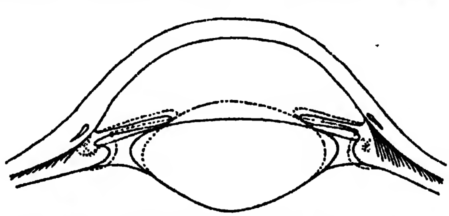


FIG. 24·6.—Section through Anterior Part of a Human Eye.

termed ***emmetropic***; if it cannot see the distant object clearly the eye is said to be ***ametropic.*** In order that a person who has emmetropic eyes may see near objects clearly, he must be able to produce a distinct image of them on the retina. This entails a diminution in the focal length of the lens and this is brought about by an increase in the curvature of the anterior surface of the lens; the posterior surface does not assist in producing the effect required [cf. Fig. 24·6]. The ciliary muscle is the motive power which causes this change in curvature to take place and the action of the ciliary muscle is to draw forward the choroid so that the tension in the suspensory ligament is diminished. At the same time, as is readily observed by viewing an eye, the iris contracts so that the pupil becomes smaller. This increase in the refractive power of the crystalline lens is referred to as ***accommodation.*** Very early writers on this subject believed that an eye was able to accommodate the images of near and of distant objects by shifting the retina.

The degree of accommodation varies with the age of the individual; it becomes less with advancing years and is attributed to a gradual hardening of the eye lens. For all persons there exists a point in front of which it is not possible for an object to be seen clearly—this is called the ***near-point*** or ***punctum proximum;*** the position beyond which a distant object cannot be seen is termed the ***far-point*** or ***punctum remotum.*** In Fig. 24·7 these are denoted by P_P and P_R, respectively. For normal eyes the far-point is at infinity, while the near-point is about 20 cm. away (for infants it is sometimes as low as 7 cm.).

Experiment.—To locate the near-point.—A lens of about 10 cm. focal length is held very near to the eye and a small object is moved until its image is clearest (as in the correct use of a magnifying glass, p. 501). From the known relative positions of the lens and object, and the known focal length, the position of the image is calculated. This is the distance of the near-point from the eye: it is known as the ***least distance of distinct vision.***

Some Defects of Vision.—An eye which is capable of producing a clear image of a distant object, i.e. one which brings parallel incident light to an exact focus on the retina, and which possesses the average power of accommodation for the particular age of the person concerned, is said to be ***emmetropic*** [cf. Fig. 24·7 (*a*)]. When such an eye views an object close at hand, we have seen that the ciliary muscle causes the suspensory ligament somewhat to relax so that the anterior surface of the crystalline lens becomes more curved. For normal persons at an early age the images of objects from 7 cm. to infinity may be focused sharply on the retina, but the range of accommodation progressively becomes less with advancing years. This defect of vision is term ***presbyopia*** and is

caused by a gradual hardening of the lens. It becomes a source of inconvenience when the near-point has receded beyond a com-

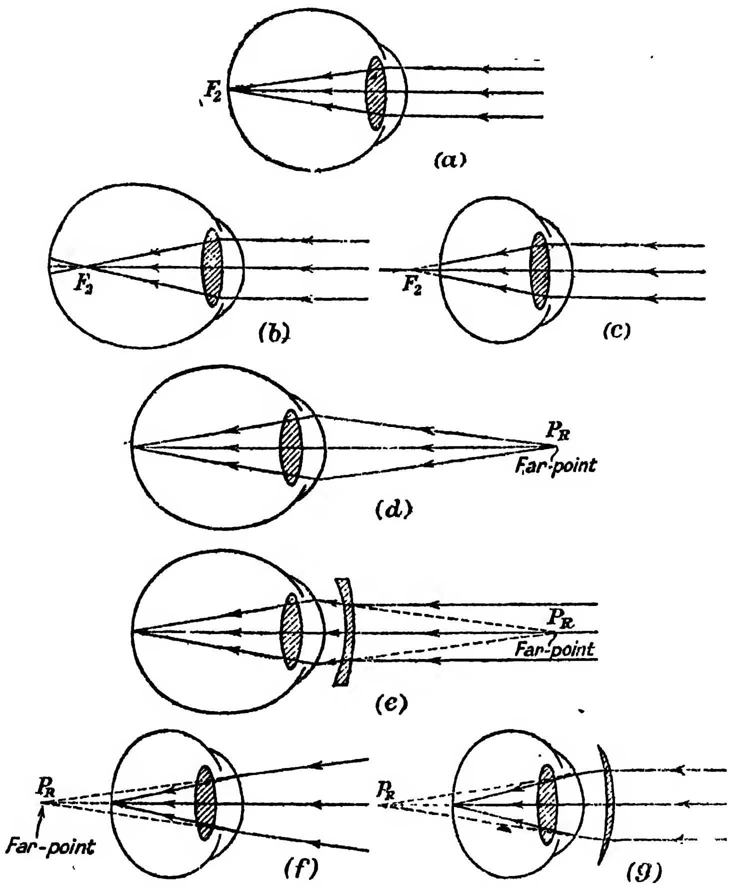


FIG. 24·7.—Some Defects of Vision and the Methods used to Correct Them.

(a) An emmetropic or normal eye: far-point at infinity.
(b) An ametropic eye—axial length greater than normal: myopia or short-sight.
(c) An ametropic eye—axial length less than normal: hypermetropia or long-sight.
(d) A short-sighted or myopic eye cannot see clearly objects beyond the far-point.
(e) Correction of (d) with the aid of a diverging lens.
(f) A long-sighted or hypermetropic eye has a virtual far-point and when such an eye is at rest rays must be directed towards this point for a sharp retinal image to be formed.
(g) Correction of (f) with the aid of a converging lens.

N.B.—The diagrams assume that refraction occurs entirely at the anterior surface of the cornea: the principal refraction occurs here, but there is also some refraction by the lens.

fortable reading distance. This defect is most troublesome in a weak or artificial light, for the aperture must be extended, with consequent loss in definition, in order to let in sufficient light in

such instances. Low-power converging lenses are prescribed to remedy this defect of vision: they behave as a simple magnifying-lens [cf. p. 500] and form a virtual image of the object at a distance from the lens greater than the object distance itself, i.e. the object may be closer actually to the eye than its own near-point.

An eye whose far-point is not at infinity is said to be ***ametropic.*** The two most important forms of ametropic eyes are those in which the axial length, i.e. the distance from the cornea to the retina, is either excessive or defective. The state of the eye in which the axis is increased beyond its normal length is referred to by the terms ***myopia*** or ***hypometropia***; the condition in which the axis is less than its normal value is called ***hypermetropia.***

In ***myopia*** (or short-sight) the image of a distant object is formed at a distance in front of the retina; in ***hypermetropia*** (or long-sight) the focus for parallel light is beyond the retina.

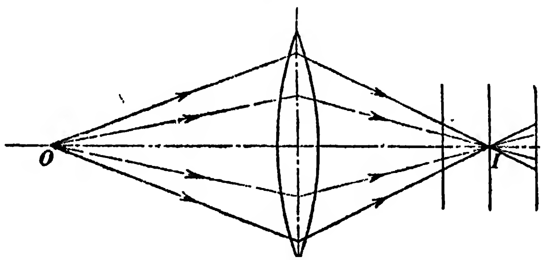


FIG. 24·8.—To Illustrate the Reduction in Area of the Circle of Diffusion when the Aperture of a Lens is Reduced.

In both these defects the image on the retina is diffuse—every point in the object has a corresponding circle of illumination on the retina—this is called the ***circle of diffusion,*** its formation being shown in Fig. 24·7 (*b*) and (*c*). These defects can be corrected by a suitable choice of spectacles; the defect is easily detected in persons whose eyes are ametropic even if glasses are not worn, for it is noticed that such persons tend to make the pupil of the eye contract, i.e. the aperture of the lens is reduced. When the aperture is so diminished the diffusion circles become less, i.e. the image is more distinct. These conditions are illustrated in Fig. 24·8.

Another very common defect is that of ***astigmatism.*** An eye is said to be astigmatic when it has different refractive powers in different planes; these differences can very frequently be attributed to anomalies in the curvature of the cornea. These irregularities cause an image produced by the rays in one plane to be brought to a focus sooner than those which are in another

plane, e.g. it may be possible to see clearly the vertical lines in a diagram when the horizontal ones are blurred or even not visible. This particular defect is corrected by the use of cylindrical lenses.

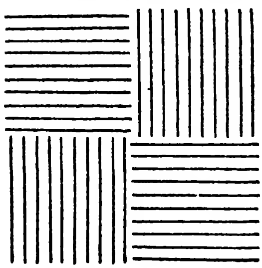


FIG. 24·9.—Simple Test for Astigmatism.

A very simple test for astigmatism is to view the diagram shown in Fig. 24·9. An astigmatic eye which is focused so that the vertical lines are clear, fails to see the horizontal lines distinctly.

Vision through a Lens. — (*a*) ***Myopia or short-sight.***—The image of a distant object is, in the case of a short-sighted person, brought to a focus in front of the retina. In order to produce a clear image on the retina the focal length of the crystalline lens must be increased. The ciliary muscles having failed to do this a supplementary lens must be placed in front of the eye. This lens must be diverging because, as indicated in Fig. 24·7 (*d*) and (*e*), the far-point of a myopic eye is at a finite distance in front of the eye, and this point must coincide with the second principal focus of the diverging lens if parallel rays are to be focused on the retina.

Example.—A person finds that his maximum distance of distinct vision is 150 cm. What spectacles will he require in order to view a distant scene ?

The lens required is such that when it is used close to the eye the image of a distant scene must be 150 cm. in front of the lens. The second focal point of the lens is therefore 150 cm. in front of the lens, i.e. the lens is a diverging one, its focal length being numerically 150 cm. [cf. Fig. 24·7 (*d*) and (*e*)].

(*b*) ***Hypermetropia or long-sight.***—Hypermetropia is the more common defect of vision, most 'normal' eyes being slightly hypermetropic. When rays from a distant object enter the eyes of a person suffering from this defect, they are refracted so that they tend to form a clear image behind the retina. Moreover, since slightly converging rays entering a hypermetropic eye must possess the correct amount of convergency in order for them to be focused on the retina, it follows that the far-point for a hypermetropic eye must be behind the retina, i.e. it is virtual. Hence, if parallel rays are to be focused on the retina of a hypermetropic eye, they must be given the required amount of convergency before entering the eye: this is effected by means of a converging lens, generally a meniscus one, for this increases the field of view.

Example.—The far-point for a certain hypermetropic eye without accommodation is 30 cm. behind the cornea. What is the focal length of the lens, held close to the eye, which will render an object 50 cm.

away clearly visible ? If the lens is moved 1 mm. from the eye, how must the object be moved in order that it still shall be clearly visible ?

The lens required must be able to converge rays of light from an object 50 cm. away so that in the absence of the eye a real image would be formed 30 cm. from the lens. Its focal length, f, is therefore given by

$$\frac{1}{f} = \frac{1}{v} - \frac{1}{u} = -\frac{1}{30} - \frac{1}{50}. \qquad \therefore f = -18{\cdot}75 \text{ cm.}$$

A converging lens of focal length 18·75 cm. is therefore required.

When the lens is shifted 0·1 cm. away from the eye, let the object distance be U, measured from the new position of the lens. Since the image distance is now − 30·1 cm.,

$$-\frac{1}{30{\cdot}1} - \frac{1}{U} = -\frac{1}{18{\cdot}75}, \quad \text{or} \quad U = 49{\cdot}71 \text{ cm.}$$

Hence the object must be 49·81 cm. from the initial position of the lens, i.e. the object must be brought 0·19 cm. nearer to the lens.

(c) ***Presbyopia.***—It has been stated already that this defect of vision is due to a decrease in the power of accommodation of the eye, i.e. the eye-lens cannot be controlled so that its focal length

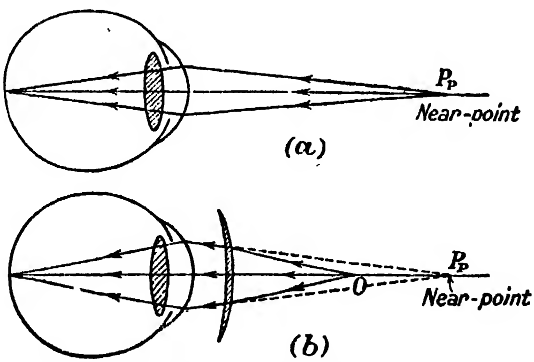


FIG. 24·10.—Presbyopia and how to Correct it.

becomes sufficiently short for close objects to be seen clearly. A supplementary converging lens, whose focal length may be calculated when the position of the near-point of the defective eye is known, is therefore required [cf. Fig. 24·10 (a) and (b)].

Example.—**A person finds that his near-point is 75 cm. away. He wishes to read print at a distance of 25 cm. from the eye. What lens, held close to the eye, is required ?**

Let f be the focal length of the auxiliary lens. Then it must be such that when the object is 25 cm. from it, the image formed by refraction through it is virtual and 75 cm. away, i.e. $v = +75$ cm. when $u = +25$ cm.

$$\therefore \frac{1}{f} = \frac{1}{75} - \frac{1}{25}, \quad \text{or} \quad f = -37{\cdot}5 \text{ cm.}$$

A converging lens of focal length 37·5 cm. is therefore required.

Some Optical Illusions.—The sketches which comprise Fig. 24·11 are examples of some familiar illusions. The straight lines in the first diagram are parallel, although they do not appear so; the fact that they are continuous and parallel can be verified by viewing them sidewards from a low point of vision. The square of the next

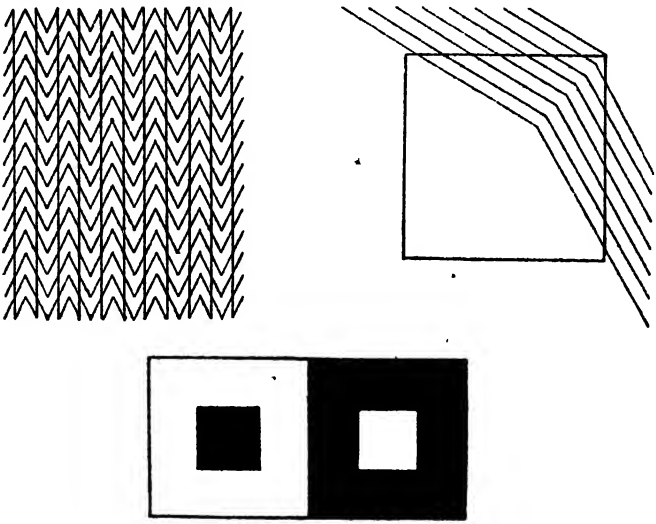


Fig. 24·11.—Some Optical Illusions.

figure has been drawn accurately, yet this is apparently not so. The white square upon the black ground looks larger than the black square upon a white ground, although the two are equal. The reason for this is that the image of a point is a small circle, so that the edges of the white regions invade the black ones. This is the so-called ***irradiation***, a phenomenon to which all the above illusions can be attributed.

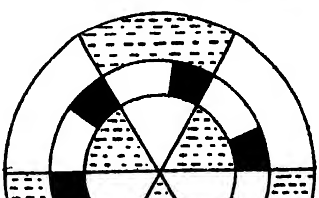


Fig. 24·12.

Retinal Fatigue.—Suppose that a disc is painted as shown in Fig. 24·12. The shaded portions are red, the rest are black and white. What will happen when such a disc is rotated in a plane about its centre? One would imagine that there would be an inner and outer portion which would appear pink, the region in between being grey—a mixture of white and black. Such is only partly the case: the pink is there, but where one would have expected the grey there

is a faint bluish-green colour. The reason for this is that the eye becomes fatigued to the pink colour. All the white light from the central region does not stimulate the retina. The retina is tired of red light and only records the other colours—viz. the complementary bluish-green colour.

The Young-Helmholtz Theory of Colour Vision.—ABNEY first showed experimentally that any colour may be matched visually by adding together various amounts of the three primary colours. Let it be assumed that one half of a field of view is illuminated by a light stimulus, Q, having any desired energy distribution. This field may be matched in the other half by mixing the light from the three primaries in amounts R, G, and V respectively, i.e.

$$Q = R + G + V.$$

If one of these quantities is negative, that primary must be added to that part of the field which is illuminated by the unknown stimulus.

The Young-Helmholtz theory of vision assumes that the eye contains three independent nerve sets, each being a selective detector of light energy. When more than one set of detectors is excited, a mixed sensation is produced, its character depending on the degree to which the individual sets have been stimulated. According to this theory it is assumed that each set of detectors, the red for example, transmits only the sensation of red to the brain, independently of the manner in which it is excited, i.e. light of a colour which is not red affects the red detectors to some extent, and the impression on the brain is that of red light.

Although this theory is a useful one in helping us to understand the mechanism of a light sensation, there are some serious objections to it. One of these is that there is no anatomical evidence for the existence of three sets of nerves in the retina.

Hypochromatic Vision or Colour-Blindness.—The Young-Helmholtz theory of colour vision accounts for the colour sensations of colour-blind persons. These are people in whom one of the sets of selective detectors is not operative, e.g. the red sensation may be missing. As a rule such people confuse red and green objects. The employment of such a person as a driver of a railway train would result, sooner or later, in an accident. Occasionally, two sets are inoperative: when this occurs the sensation produced is very much like the black and white rendering of a coloured object in a photographic print. These people match every colour with some shade of grey, for the only sensation they perceive is that which normal eyes interpret as white. Hence, to them, colours only differ from each other and from white in the degree of brightness,

Colour-blindness was formerly termed *Daltonism*, since DALTON suffered from this malady. He was unaware of this defect in his vision until 1792 when he noticed that a pink geranium was pink by candle-light, but sky-blue by day. He examined the spectrum of white light and found that the image termed red by others was little more than a "shadow or a defect of light." Orange-yellow and green seemed one colour, while there was a pronounced difference between blue and green. Dalton said that a florid complexion looked blackish-blue on a white ground—persons with normal vision may obtain an idea of this effect by observing people in the light from a mercury vapour lamp. Dalton also maintained that a laurel leaf was a good match to a stick of sealing-wax. The following story about Dalton is at least amusing. He, being a Quaker, objected to wearing any material which was scarlet in colour. However, he wore a doctor's robe (scarlet) for several days without realizing the astonishment it caused to others.

The Use of Ultra-Violet Light in Therapeutics.—The spectrum which is visible to the eye is only a small portion of the complete spectrum of æthereal waves. The region beyond the violet end of the spectrum is the ultra-violet region and the wave-lengths here are shorter than in the visible spectrum. The natural source of ultra-violet radiation is the sun, but the amount of ultra-violet light which reaches any particular place depends, amongst other things, upon the altitude of the sun and the amount of atmospheric pollution. The greater the altitude the less the distance in air traversed by the sun's rays so that the rays are less absorbed. The intensity of ultra-violet light is a maximum about 1 p.m. on a clear day. In the immediate neighbourhood of a large industrial city the amount of such radiation present in the rays which finally reach the earth's surface is practically zero. Recent research has shown that ultra-violet radiation is essential for the well-being of the community so that, in places where sufficient ultra-violet radiation is not to be obtained from the sun, artificial sources must be used. Chief among such sources of this so-called artificial sunlight are the mercury vapour lamp and the tungsten arc. The mercury vapour lamp [cf. p. 471] consists of a silica vessel containing mercury and its vapour only. When a suitable potential is applied to the tube, a brilliant green light is seen and much ultra-violet light is emitted. In the tungsten arc lamp an electric arc is formed between tungsten poles and is a very powerful source of such radiation; in fact the patient must wear dark glasses in order to protect the eyes. If the eyes are not so protected, permanent blindness may follow.

In using these sources of ultra-violet rays in the home persons must be careful to guard against an over-dose. One of the worst of all the 'light' diseases which may be produced is *Xeroderma*

pigmentosum; coloured spots begin to appear on the skin, and, in early adolescence, may prove fatal. In fact, ultra-violet lamps should not be used too liberally, and it is better to obtain medical advice.

Under suitable restrictions ultra-violet rays have proved themselves to be very beneficial. When an organic compound called ***ergosterol*** is exposed to ultra-violet radiation, vitamin D is produced. The ergosterol loses its crystalline form and becomes resinous. This vitamin is essential if rickets are to be cured, and it has been shown recently that the decay of teeth (***caries***) is largely due to a deficiency of this vitamin in early childhood. Recent work has also shown that the stamina and milk of cows are improved when they are subjected to this so-called artificial sunlight or ultra-violet radiation. It has also been found that fowl lay better and that the eggs produce more healthy chickens after such treatment.

EXAMPLES XXIV

1.—A magnifying glass is held 3·6 in. in front of a newspaper and the print appears to be 3 times as large. What is the focal length of the lens?

2.—A person can see distinctly at a distance of 4 ft. What lens must be used in order for him to see a person 20 ft. away clearly?

3.—A man cannot see distinctly unless the object is 40 in. away. He holds a book 15 in. from his eyes when reading. What sort of lens must be used? What is the power of this lens?

4.—A man can see distinctly at a distance of 27·5 in. What lenses are necessary so that he may read a book 16·2 in. away?

5.—Explain the use of a converging lens as a magnifying glass. How is its magnifying power defined? A magnifying glass of 5 cm. focal length is used by a person whose least distance of distinct vision is 25 cm. Calculate the best position of the object, and the magnifying power of the lens, when the person holds it close to his eye.

6.—Describe the optical system of the eye, and explain how three common forms of defective vision may be remedied by means of spectacles.

CHAPTER XXV

THE ELEMENTARY THEORY OF SIMPLE OPTICAL INSTRUMENTS

The Visual Angle and Visual Acuity.—The apparent size of an object viewed by an unaided eye depends solely upon the size of the retinal image, and this is determined by the angle which the object subtends at the eye. It must be remembered that the eye-lens cannot be treated as a single thin lens, so that for the purpose of constructing the size of the image formed on the retina it is usual to regard the point H, Fig. 25·1, as the 'centre of the

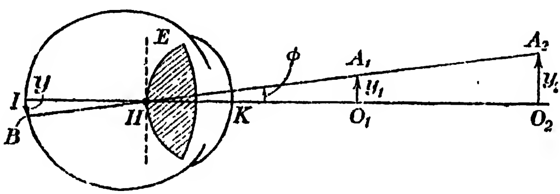


FIG. 25·1.

lens'. The point H is located where the posterior surface of the eye-lens cuts the axis of the eye-ball.

Fig. 25·1 shows how to determine the actual size of an object O_1A_1, of height y_1, when viewed by an unaided eye. It is only necessary to join by straight lines the points O_1 and A_1 to H and produce them to cut the retina in I and B respectively to give the retinal image IB of height y. Now the angle ϕ, which the object O_1A_1 (and also the image IB) subtend at H, is called the ***visual angle,*** and since the distance HK, where K is the point at which the cornea is cut by the axis of the eye, is small compared with the least distance of distinct vision for an unaided normal eye, the visual angle may be taken as the angle which an object subtends at a point on the exterior surface of the cornea. If D is the ***least distance of distinct vision,*** i.e. the distance of the near-point from the eye, then the greatest useful value of the visual angle for an object of height y_1 is $\frac{y_1}{D}$. If the object is brought nearer to the

eye than is the near-point the angle subtended by the object, i.e. the visual angle, is increased, but the image becomes indistinct.

The diagram shows that the retinal images of an object O_2A_2, of height y_2, and an object O_1A_1 of height y_1 and arranged as in Fig. 25·1, are identical in size since the visual angles are equal.

In connexion with retinal images it must be noted that two points in an object (or two point objects) cannot be distinguished if the images are too close together. The reason for this is that the structural arrangement of the cells in the retina resembles a mosaic pattern, each cell or 'tile in the pattern' having a finite size, so that separate images are not discerned unless each falls on a different cell. When the images are formed on the most sensitive part of the retina (i.e. the fovea centralis) it is not possible for the images of two point objects to be separated unless their angular separation exceeds one minute. This limiting angle is known as the ***visual acuity*** of the eye.

Magnifying Power or Angular Magnification.—The visual angle, or apparent size of an object, increases as the object is brought nearer to the eye, but it cannot be brought nearer than about 25 cm., the least distance of distinct vision for an unaided normal eye, without appreciable loss of definition. At the least distance of distinct vision the object is seen most clearly by an unaided eye, but it must be remembered that, even so, any details in the object whose angular separation is less than about one minute will not be seen. Similarly, when we view two close stars, i.e. two stars whose angular separation tends to zero, the eye will not distinguish them as separate entities if the visual angle for the pair of stars is less than one minute. In such cases the eye is said to fail to resolve the objects. To overcome these limitations and so extend the range of clear vision optical instruments have been designed. In these instruments lenses are arranged so that when an object is viewed through them a virtual image subtending an angle greater than the visual angle of the object (or the details in it) is obtained: the eye observes this image instead of the direct object.

Before proceeding further it is necessary to explain what is meant by the ***magnifying power***, or ***angular magnification***, M, of an optical instrument. It is defined by the equation

$$M = \frac{\text{Angle subtended by image seen through the instrument}}{\text{Angle subtended by object seen by the naked eye, under suitable conditions}}.$$

Now the angle subtended at the eye by the image depends somewhat, but not to any large extent, on the adjustment of the particular instrument in use, but the angle subtended by the object

depends altogether on its position so that this must be specified carefully. For microscopes the distance of the eye from the object when calculating the visual angle for the latter is taken as the nearest distance of distinct vision, usually a conventional 25 cm. which, in theoretical work, is denoted by D. With telescopes, the object being *in situ*, the position of the observer is not of much consequence. On the other hand, for short-range telescopes the object is *in situ* and the observer can be either *in situ*, or, as will be discovered later, if his eye is considered to be at the objective of the telescope the expression for the magnifying power of a short-range telescope is considerably simplified [cf. p. 515].

The Magnifying Glass, Reading Lens, or Simple Microscope.—Let OA, Fig. 25·2, be a small object within the first focal distance of a converging lens L, whose principal foci are F_1 and F_2 respectively. To locate the image we consider a ray AH parallel to the principal axis of the lens; after refraction, this proceeds in the direction HF_2. Since the ray AC passing through the centre of the lens does so without deviation, it follows that the image of A must be B, the point at which F_2H and CA meet when produced.

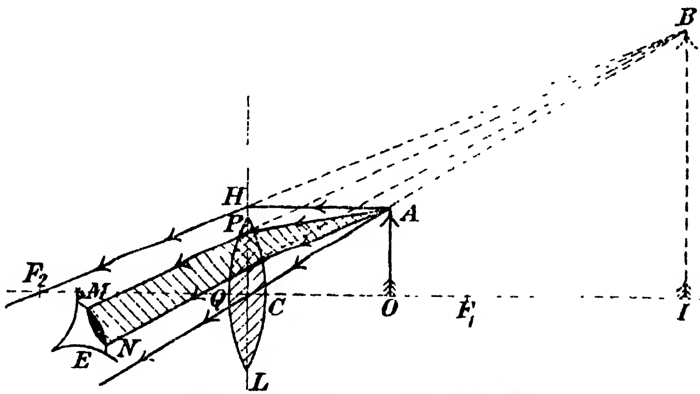


FIG. 25·2.—A Simple Microscope.

The image of OA is therefore IB: it is erect, magnified, and virtual. If an eye is placed at E and this image viewed, all the rays incident upon the lens do not necessarily enter the pupil of the eye after refraction by the lens. To obtain the confines of the rays which enter the eye we first join M and N, the extremities of the pupil, to B. If these intersect the lens in P and Q, then PM and QN are the rays which, having passed through the lens, enter the eye. Since, B is the image of A it follows that AP and AQ must be the rays which travel from A to the eye. Similarly, by joining M and N to I the confines of the rays proceeding from O to the eye are obtained.

The Magnifying Power or Angular Magnification of a Simple Microscope.—In dealing with the principle of a simple microscope only the manner in which the image is formed has been discussed. It is now necessary to consider the position of the image. Usually, when observing the image of a near object

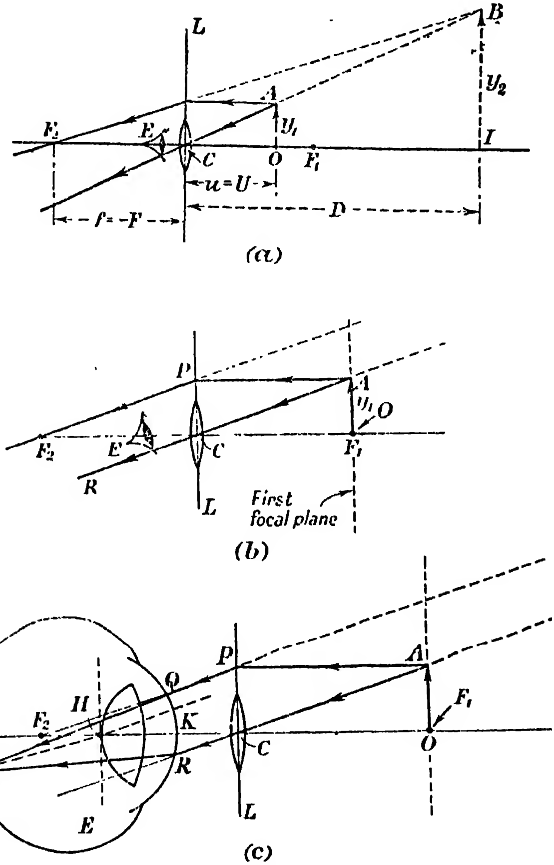


Fig. 25·3.

with respect to the lens, this is adjusted almost self-consciously by the observer so that the image is produced at his least distance of distinct vision. To use a simple microscope correctly one should therefore proceed as follows. The eye is placed as close as possible to the lens, the object is held near to the lens and then the relative distance of lens and object increased until the image is distinct. The reason for this procedure is that when it is adopted the field

of view is a maximum and, as the sequel shows, the magnifying power is then as great as possible.

When, however, a single converging lens (or its counterpart in an actual instrument) is used as an eye-piece of a telescope, especially if this forms part of a spectrometer, then both eyes of the observer should remain open, one to view a distance object unaided and the other to look through the lens at the cross-wires. These should be made to lie in the first focal plane of the lens so that the image seen through it is formed at infinity. One advantage of this procedure is that long-sighted persons can use the instrument when it has been adjusted by a person with normal vision.

Of course the image seen through a simple microscope can be located in any position between the near point and infinity, but it will be found that the magnifying power of a simple microscope does not vary very much between the extreme values it may assume. Consequently only the values of these two extremes will be calculated.

(*a*) *The image seen through a simple microscope is formed at the nearest distance of distinct vision.* To simplify the discussion it will be assumed, at first, that the observer's eye is close to the lens. Then let OA, Fig. 25·3 (*a*), be a small object viewed through a thin converging lens of focal length $f = -F$ (say), the image being IB and it is supposed to be formed at the least distance of distinct vision, i.e. $CI = D$. [Details of how to locate the image IB are not repeated here.] Let y_1 and y_2 be the heights of the object OA and its image IB respectively, and let $u = U$ be the distance of the object from the lens. Then if an eye is set at E, close to the lens L, the angular magnification or magnifying power, denoted in this instance by $M_{v=D}$, is given by

$$M_{v=D} = \frac{\text{Angle subtended at E by the image}}{\text{Angle subtended by object when at a distance D from the observer}}$$

$$= \frac{\dfrac{|y_2|}{|D|}}{\dfrac{|y_1|}{|D|}} = \frac{|y_2|}{|y_1|} = \frac{D}{U},$$

since the $\triangle$'s BC and ACO are similar. Thus in this instrument, as used under the conditions postulated, the angular magnification is identical with the longitudinal magnification [cf. p. 430].

Since F is the numerical value of the focal length of the lens

$$\frac{1}{D} - \frac{1}{U} = -\frac{1}{F}, \quad \text{or} \quad \frac{D}{U} = 1 + \frac{D}{F},$$

$$\therefore\ M_{v=D} = 1 + \frac{D}{F}.$$

(*b*) *The image seen through a simple microscope is formed at infinity.* —When this occurs the object, OA, must lie in the first focal plane of the lens. Fig. 25·3 (*b*) shows how to determine the paths of the rays which come from the point A in the object and emerge after refraction through the lens so that an eye at E sees an image of OA at infinity. In this instance let $M_{v \to \infty}$ denote the magnifying power of the instrument. Then, by definition,

$$M_{v \to \infty} = \frac{\text{Angle subtended at E by the image}}{\text{Angle subtended by object when situated at a distance D from the observer}}$$

$$= \frac{\frac{y_1}{F}}{\frac{y_1}{D}},$$

since both image and object subtend $\widehat{ACF_1}$ at E when E is close to the lens.

$$\therefore\ M_{v \to \infty} = \frac{D}{F}.$$

Thus for a converging lens of focal length 5 cm., i.e. $f = -5$ cm. or $F = 5$ cm., $M_{v=D} = 1 + \frac{25}{5} = 6$, and $M_{v \to \infty} = 5$, so that the difference between the limits for the magnifying power is less than twenty per cent. of the greatest value $M_{v=D}$. With lenses of shorter focal length, i.e. higher magnifying power, the percentage difference is less.

In connexion with this problem of image formation through a simple microscope when the image is at infinity, it is instructive to show how the size of the retinal image may be found. Thus in Fig. 25·3 (c), let PQ and CR be two of the rays in the pencil of light which had its origin at A but which did not become a parallel beam until after refraction through the lens had occurred. Let H [cf. p. 498] be the 'effective centre' of the eye E. Then if through H a straight line HB is drawn parallel to PQ (or to CR), then all light from A will be brought to a focus at B; IB is the retinal image required and APQB and ACRB are the paths of two rays from A to B.

More about the Magnifying Power of a Simple Microscope. —Hitherto it has been assumed that the observer's eye has been very close to the lens when the latter is used as a simple microscope. Now, in practice, this condition cannot always be fulfilled, and with lenses of short focal length the actual distance of the eye from the lens may be comparable with the focal length of the lens itself. Let us therefore see how the angular magnification depends upon the distance of the eye from the lens.

In this connexion let us suppose that an eye E is situated on the principal axis of a simple microscope, Fig. 25·4, and let the numerical value of the distance of the eye from the lens be A. Let IB be the image seen through the lens of an object OA and let this image be produced at the least distance of distinct vision from E. Then the image is at a distance (D — A) from the centre

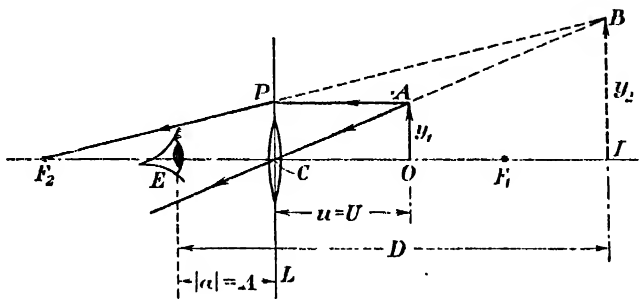


Fig. 25·4.

of the lens. Let $M_{v=(D-A)}$ be the angular magnification under these circumstances. We have

$$M_{v=(D-A)} = \frac{\text{Angular size of image}}{\text{Angular size of object when at the near-point of the observer}}$$

$$= \frac{|y_2|}{|D|} \div \frac{|y_1|}{|D|} = \frac{|y_2|}{|y_1|} \qquad \text{[N.B. } |D| = D\text{]}.$$

$$= \frac{D-A}{U} \qquad \text{[By similar } \triangle\text{'s ACO and BCI]}.$$

But
$$\frac{1}{D-A} - \frac{1}{U} = \frac{1}{F},$$

so that

$$M_{v=(D-A)} = (D-A).\frac{(F+D-A)}{F(D-A)} = \frac{F+D-A}{F}.$$

When $A \to 0$ this expression reduces to the value already obtained for $M_{v=D}$. When the image is formed at a distance, whose numerical value is V, from the lens, we have

$$M_{v=V} = \frac{|y_2|}{V+A} \div \frac{|y_1|}{D} = \frac{|y_2|}{|y_1|}.\frac{D}{V+A}$$
$$= \frac{V}{U}.\frac{D}{V+A}.$$

But $\frac{1}{V} - \frac{1}{U} = -\frac{1}{F}$, so that $V = \frac{UF}{F-U}$ and therefore

$$M_{v=V} = \frac{F}{F-U} \cdot \frac{D}{\frac{UF}{F-U} + A} = \frac{FD}{U(F-A) + AF}.$$

As a check it should be noted that when $U = F$ so that $V \to \infty$, the expression just obtained for the angular magnification becomes equal to $\frac{D}{F}$, which is the value previously obtained for $M_{v \to \infty}$. The present discussion reveals the fact that when $v \to \infty$ the angular magnification is independent of the position of the eye of the observer with respect to the lens.

On the Achromatic Nature of the Image seen through a Simple Microscope.—On account of the chromatic aberration which has been shown to exist when a real image is produced by a thin converging lens it is natural to expect that chromatic effects would spoil the definition of the image seen through a simple microscope. Experimentally it is found that the centre of the field of view is remarkably free from chromatic defects : it is only in the peripheral region of the field that the chromatic defects are troublesome. Referring to Fig. 25·2, it is seen that the angle which the image subtends at the centre of the lens is always $A\hat{C}O$, so that the apparent size of any one coloured image is independent of colour : the differently coloured images are apparently superimposed so long as we restrict ourselves to the region of paraxial rays, and it must be remembered that nearly all our discussions have been limited to the region of such rays. The coloured effects which appear in the outer regions of the field of view are due to differences in the spherical aberration for light of different wavelengths.

If cross-wires illuminated by white light are observed through a simple microscope, the central portion of the image is remarkably free from colour, but if the same lens is used to produce a real image of the wires on a screen, chromatic defects are much in evidence.

In connexion with the achromatism of a simple microscope it must be noted that since all the coloured images have the same apparent size, achromatism is obtained for all colours : when an achromatic lens combination is used to produce a real image the achromatism is only obtained with respect to two colours.

Experimental determination of the Magnifying Power of a Simple Microscope.—(*a*) *The image is formed at the least distance of distinct vision.*

(i) First let it be assumed that the observer is close to the lens. Then if F is the numerical value of the focal length of the lens and D the least distance of distinct vision, it has been shown, [cf. p. 502], that

$$M_{v=D} = 1 + \frac{D}{F}, \qquad \left[\text{or } 1 + \frac{D}{|f|}\right].$$

Hence to measure the angular magnification due to a converging lens it is only necessary to measure the focal length of the lens and then determine the distance at which a small object must be situated for the image to be formed at the least distance of distinct vision: the value of D is calculated from the known value of the focal length of the lens and the object distance.

In practice the above method is objectionable since all observers do not obtain the image in the same position relative to the lens. The following method is preferable. The lens L, Fig. 25·5 (*a*), is mounted in a vertical stand and a cover slip, S, attached to it with the aid of soft wax so that its plane is inclined at about 45° to the principal axis of the lens. A piece of graph paper, BG, is then placed 25 cm. away from S as shown. [We neglect the distances of the observer and of the cover slip S from the lens.] The paper BG should be well illuminated and every fifth millimetre division inked over—one of these inked lines, near the centre, should be red, the rest black. The distance between any two such lines constitutes a 'division.' An observer at E will then see an image I_1 of the scale BG—it is formed by reflexion at S and is not magnified.

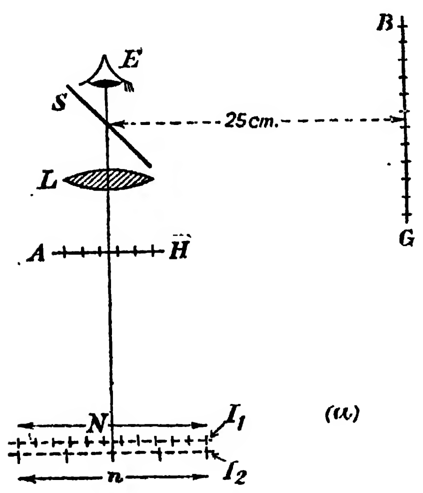


FIG. 25·5 (*a*).

Another piece of graph paper, AH, similar to BG but without the line in red ink, is then placed behind the lens and its position adjusted until the image of a line upon it coincides with the image of the red line formed by reflexion at S. The two images are then coplanar, i.e. the image I_2 formed by refraction through the lens is 25 cm. away from it. If N is the number of divisions in I_1 which appear to coincide with n divisions in I_2, the required angular

magnification is $\frac{N}{n}$. The value thus obtained should be compared with the calculated value $\left(1 + \frac{25}{F}\right)$.

(ii) In the second place, let the cover slip be at a distance $|a| = A$ from the lens and assume that an eye may be placed close to this cover slip. The experimental arrangement is shown in Fig. 25·5 (*b*)

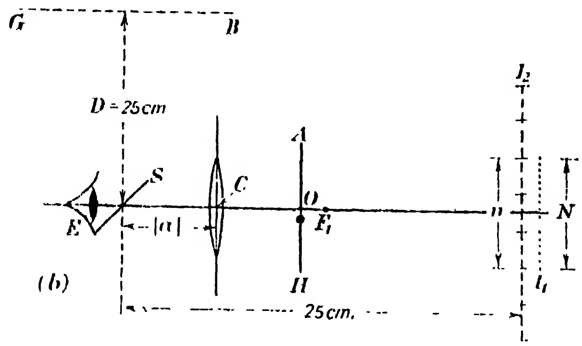


Fig. 25·5 (*b*).

and no further details should be required. The magnifying power is again $\frac{N}{n}$ and this should be equal to $M_{v=25-A} = \frac{F + 25 - A}{F}$.

(*b*) *The image is formed at infinity.*—The converging lens L_1, which is used as a simple microscope, is adjusted so that an object OA, Fig. 25·6, lies in its first focal plane.

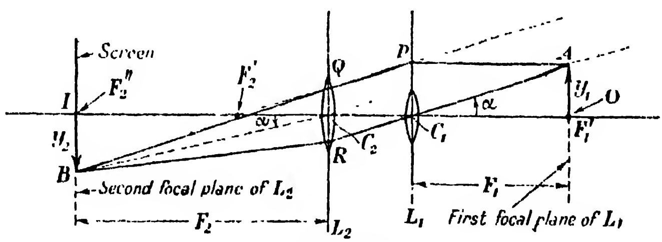


Fig. 25·6.

Let the object OA consist of two parallel wires stretched across an aperture which is suitably illuminated: let y_1 be the distance apart of these wires. Then the angle, α, which these wires subtend at C_1, the optical centre of the lens, is given by $\alpha = \frac{|y_1|}{F_1}$, where F_1 is the numerical value of the focal length of this lens. [A plane mirror may be used to set the lens L_1 in the correct position with

respect to the object—cf. the autocollimating method, p. 439.] In order to measure the angle α experimentally a converging lens, L_2, of known focal length $f_2(= - F_2)$, is placed coaxially behind L_1 an image, IB, of the illuminated wires focused on a translucent screen situated in the second focal plane of L_2 is obtained. Then since all rays from A emerge from L_1 as a beam of light parallel to the direction AC_1 the point B in the image on the screen corresponding to the point A in the object is obtained by drawing through C_2, the optical centre of L_2, a straight line C_2B parallel to AC_1 to cut the second focal plane of L_2 in B. If y_2 is the height of the image IB, the angle α is given by

$$\alpha = \frac{|y_2|}{F_2} = \frac{Y_2}{F_2}, \qquad [\text{if } Y_2 = |y_2|],$$

so that

$$M_{v \to \infty} = \frac{\alpha}{\dfrac{|y_1|}{|D|}} = \frac{Y_2}{F_2} \cdot \frac{D}{Y_1}, \qquad [\text{if } Y_1 = |y_1|].$$

The value so obtained should be compared with $\frac{D}{F_1}$, which is the magnifying power of the lens L_1 under the conditions postulated.

The Compound Microscope.—The magnifying power of a thin converging lens has been shown in the preceding pages to have its maximum value $1 + \frac{D}{F}$ when the image is formed at the least distance of distinct vision for an eye placed close to the lens. The magnifying power may be increased by decreasing the numerical value of the focal length of the lens, but such lenses are difficult to grind accurately. Moreover, spherical aberration is increased and the images seen through such lenses are badly distorted. Hence in order to obtain high magnifying powers resort must be had to other devices: an example is found in the compound microscope.

One of the first of such instruments was made by GALILEO in about 1610 soon after he had invented his telescope. In its simplest form a compound microscope consists essentially of two thin converging lenses arranged coaxially; the first lens, known as the ***objective*** or ***object-glass***, is of short focal length and is used to produce a real image of the object, while the second lens, or ***eye-piece***, is used to produce a virtual image of the image produced by the objective.[1]

[1] If Galileo actually followed the design of his telescope and used a diverging lens for the eye-piece of his microscope, the field of view would be very small. By 1624, Drebbel in England, and Metries in Holland, had produced compound microscopes in which the eye-piece was a converging lens, i.e. the field of view was considerably enlarged.

When a compound microscope is in use the eye is almost invariably adjusted so that the final image is seen at the least distance of distinct vision, but, as with a simple microscope, there is no reason why this image should not be formed at infinity or anywhere in between these two positions. First let us consider the case when the image is formed at the least distance of distinct vision and the eye is close to the eye-piece.

Let L_1 and L_2, Fig. 25·7 (*a*), be the objective and eye-piece of a compound microscope. First let us locate the position of the image of a small object OA placed perpendicularly to the axis of the system. In the illustration, the two lenses have been drawn, but it must be clearly understood that, in accordance with our

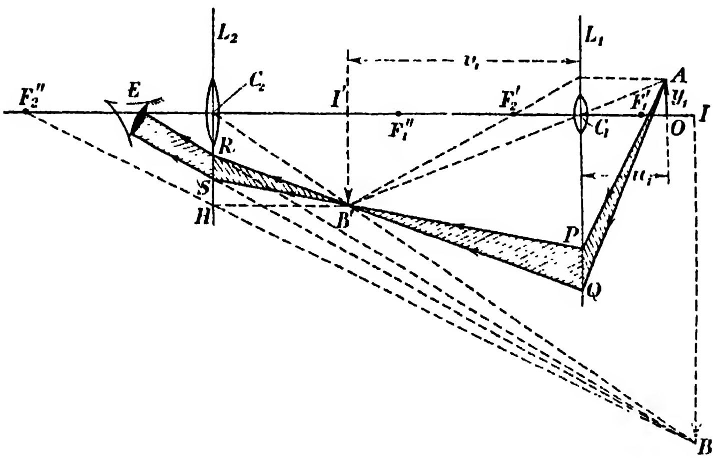


FIG. 25·7 (*a*).

usual practice in optics, the scale of the drawing in a direction at right angles to the system has been considerably enlarged to make the diagram more clear. If, therefore, some of the rays which are drawn eventually do not pass through the lens, it must be remembered that it is the lines through the centres of the lenses which are used in these constructions and that the shapes of the lenses are inserted only to remind us that such lenses are actually present.

The object is placed just beyond the first focus F_1' of the objective so that a real magnified image I′B′ is produced. It can be found in the manner indicated. The eye-lens is arranged so that I′B′ lies nearer to it than does its first focus F_1''. To determine the final image we note that a ray B′H parallel to the axis of the system is refracted by L_2 so that it passes through F_2'' the second focus of the

eye-piece. The intersection of $F_2''H$ produced and $B'C''$ the ray passing through L_2 without deviation, produced backwards gives B the image of A. By drawing BI perpendicular to the axis the final image is obtained.

To indicate the paths of the rays by which an eye sees the final image let us trace those rays which proceed from A to the eye. If additional rays have to be traced they can be obtained in the same way. The point B is joined to the limits of the pupil. Let these lines cut the plane of the lens L_2 in S and R. Joining these two points to B′ and producing them backwards to cut the plane of the objective in P and Q we obtain the required rays in the space between the two lenses. If P and Q are then joined to A the paths of the rays are obtained completely.

Secondly, let the final image be formed at infinity. For this to happen it is necessary for the intermediate image I′B′ to be formed

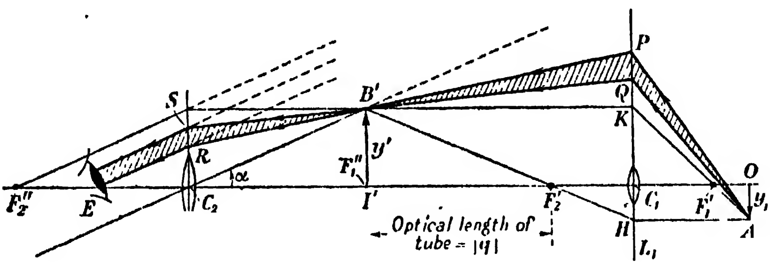


FIG. 25·7 (b).

in the first focal plane of the eye-piece. Fig. 25·7 (b) shows how to determine the paths of the rays through the instrument.

The Magnifying Power or Angular Magnification of a Compound Microscope.—(*a*) Let $M_{v=D}$ be the magnifying power when the image seen through the instrument is formed at a distance D, the least distance of distinct vision, from the eye-piece. Then, from Fig. 25·7 (*a*), we have

$$M_{v=D} = \frac{\text{Angle subtended at the eye-piece by image}}{\text{Angle subtended by object when seen at the nearest distance of distinct vision}}$$

$$= \frac{|IB|}{D} \div \frac{|OA|}{D} = \frac{|IB|}{|OA|} = \frac{|y_2||y'|}{|y'||y_1|}$$

$$= \left[1 + \frac{D}{|f_2|}\right] \frac{|v_1|}{|u_1|}, \qquad \text{[cf. p. 502]}$$

where f_2 is the focal length of the eye-piece, u_1 the distance of OA and v_1 that of the image I′B′ from the objective, and $y' = I'B'$, and $y_2 = IB$.

(*b*) Let $M_{v\to\infty}$ be the angular magnification of a compound microscope when the image seen by a normal unaided eye through the instrument is formed at infinity. Then, from Fig. 25·7 (*b*), we have

$$M_{v\to\infty} = \frac{B'C_2F_1''}{\left[\frac{OA}{D}\right]} = \frac{\frac{|y'|}{|f_2|}}{\frac{|y_1|}{|D|}}$$

$$= \frac{D}{F_2}\frac{|y'|}{|y_1|} \qquad \text{[where } F_2 = |f_2|\text{]}$$

$$= \frac{D}{F_2}\frac{|q|}{F_1'}, \qquad \text{[from similar } \triangle\text{'s } B'F_1''F_2' \text{ and } F_2'C_1H\text{]}$$

where q is the distance of the image I′B′ from the second principal focus of the objective whose focal length is numerically F_1.

It will be noted that the expression obtained for $M_{v\to\infty}$ is independent of the nearness of the observer to the eye-piece.

The Ramsden Circle.—Fig. 25·8 shows how to locate graphically in the usual way the image IB of a small object OA seen through a compound microscope. It must be remembered that all distances perpendicular to the axis of the system are considerably magnified so that the positions of the images can be located accurately. On the same scale let HK be the aperture of the

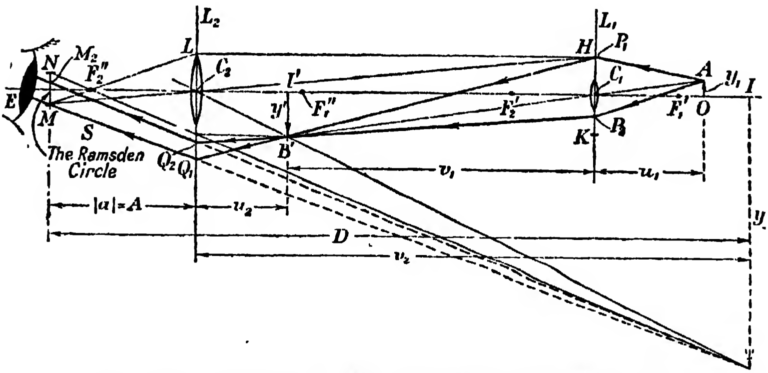


FIG. 25·8.—A Compound Microscope and its Ramsden Circle.

objective. Let MN be the image of the objective produced by the eye-piece: MN is located by means of the rays HL, which after refraction at L passes through F_2'', and HC_2, which is not deviated by the lens L_2.

Now MN is known as the ***eye-ring*** or ***Ramsden circle*** and all rays from OA which enter the microscope must emerge through MN. Hence, if the observer's eye is placed at MN and the pupil

is longer than MN, the image will be seen under the best conditions.

In Fig. 25·8 there are shown the paths of two rays, $AP_1B'Q_1M$ and $AP_2B'Q_2M_2$, through the instrument and by means of which an eye at E, close to the eye-ring MN, sees B, the image of A. It will be noticed that AP_1 just enters the objective and so just emerges through the eye-ring.

On the Magnifying Power of a Compound Microscope when the Observer's Eye is at the Eye-Ring and the Image is formed at the Least Distance of Distinct Vision.—Referring to Fig. 25·8, we have

$$M_{v_2=D-A} = \frac{|y_2|}{D} \div \frac{|y_1|}{D} = \frac{|y_2|}{|y'|}\frac{|y'|}{|y_1|} = \left[\frac{D-A}{U_2}\right]\left(\frac{F_1}{U_1-F_1}\right).$$

If L is the distance between the lenses, we have

$$\frac{1}{A}+\frac{1}{L}=\frac{1}{F_2}, \text{ and } \frac{1}{D-A}-\frac{1}{U_2}=\frac{1}{F_2}.$$

Hence, after some reduction,

$$M_{v_2=D-A} = \left[1-\frac{(A+L)(A-D)}{AL}\right]\left(\frac{F_1}{U_1-F_1}\right).$$

Experimental Determination of the Magnifying Power or Angular Magnification of a Compound Microscope.—(*a*) First let us assume that the observer is very close to the eye-piece and that the image seen through the microscope is at the least distance of distinct vision. Select two converging lenses of focal lengths 3 cm. and 5 cm. to serve as object-glass and eye-piece respectively. Mount the former, L_1, Fig. 25·9 (*a*), about 5 cm. in front of a piece of graph paper, A, every fifth millimetre division of which has been inked over, and locate the position of the real image I, produced by refraction through L_1, by placing a piece of wire held in a circular frame so that there is no parallax between the image of one of the inked lines on A and the wire. Then arrange the second lens L_2, with a cover slip C as indicated, about 4 cm. in front of the wire. The wire may then

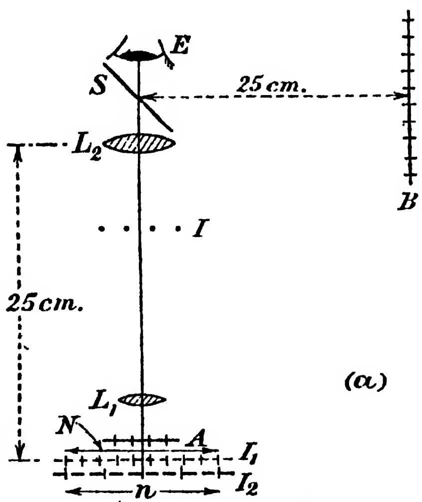


FIG. 25·9 (*a*).—Angular Magnification of a Compound Microscope.

be removed, and a piece of graph paper B, inked as described on p. 506, mounted as shown. The eye-piece L_2 should then be adjusted so that there is no parallax between the final image of one of the inked lines on A and the image of one of the lines on B, this latter image being produced by reflexion at the surface of the cover slip. The ratio of the number of divisions, N, in the image formed by reflexion at the cover slip, S, which appear to coincide with n divisions in the image produced by the microscope is the angular magnification required of the system.

(*b*) Now let the observer's eye be at the Ramsden circle appropriate to the given arrangement of lenses. The position of this circle is readily calculated when the lenses have been set in position and when the focal length of the eye-piece is known. A cover slip, S, is then placed between the eye-ring MN, Fig. 25·9 (*b*), and

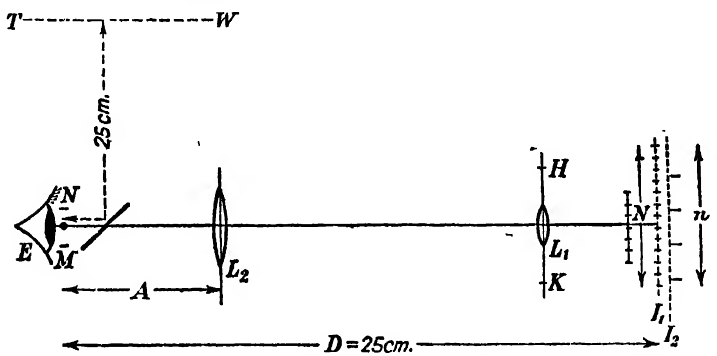


FIG. 25·9 (*b*).

the eye-piece L_2, its plane being at 45° to the axis of the microscope. A piece of graph paper, TW, is then placed at a measured distance from the axis of the instrument so that its image, formed by reflexion at the cover slip, is at a distance D (25 cm.) from MN. The ratio $\frac{N}{n}$, where the symbols have the same significance as in the preceding paragraph, is the required magnifying power. The value so obtained should be compared with that given by the formula for $M_{v_2=D-A}$ given on p. 512, and the great improvement in the field of view and appearance of the image compared with those seen when the eye is close to L_2 noted.

(*c*) When the final image is formed at infinity the angle subtended by the image must be determined as on p. 507 and the magnifying power calculated.

The Astronomical Telescope.—The essential optical features of an astronomical telescope are an objective, a converging lens of

long focal length, and an eye-piece, a converging lens of short focal length, arranged coaxially. The objective produces a real inverted image of a distant object, which is then magnified by the eye-piece. When the telescope is in ***normal*** adjustment the distance between the real image produced by the objective and the eye-piece is equal to the focal length of the latter. The final image is then at infinity. To understand the formation of this final image let us consider a parallel beam of light ACDB, Fig. 25·10, falling upon the objective whose centre is C_1 and whose second principal focus is F. After refraction by the lens, all rays in the incident beam ACDB will pass through G, the point in the second focal plane of L_1 where a secondary axis SC_1 parallel to AC cuts this plane. The extreme rays CG and DG of the refracted cone then fall upon the eye-piece and intersect the plane drawn through C_2, the optical centre of the

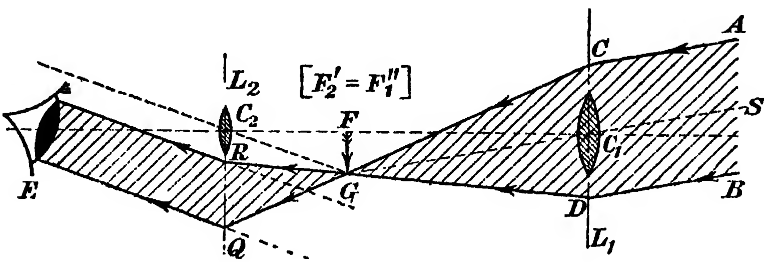


Fig. 25·10.—Principle of an Astronomical Telescope in Normal Adjustment.

lens L_2, at right angles to the principal axis of the system in R and Q. The refracted beam emerging from the eye-piece is determined by constructing straight lines through Q and R parallel to GC_2, the secondary axis of the eye-piece passing through G.

The ***magnifying power or angular magnification of a telescope*** is defined as the ratio of the angle subtended at the eye by the image to the angle subtended by the object. It is therefore given by

$$M = \frac{G\widehat{C_2}F}{F\widehat{C_1}G} = \frac{GF}{FC_2} \div \frac{GF}{C_1F} \quad \text{[since small angles may be measured by their tangents].}$$

$$= \frac{C_1F}{FC_2} = \frac{|\text{focal length of objective}|}{|\text{focal length of eye-piece}|} = \frac{|f'|}{|f''|} = \frac{F_1}{F_2}.$$

When near terrestrial objects are viewed through a telescope the latter is not in normal adjustment and the final image may be produced at any position convenient to the observer. Suppose that Fig. 25·11 (*a*) represents the positions of the two lenses L_1 and L_2

and those of the object, OA, and the final image IB for such a system. The magnifying power, M, of the system is given by

$$M = \frac{\text{Angle subtended at eye by image}}{\text{Angle subtended by object seen directly by the eye under suitable conditions}}$$

If it is stipulated that the eye shall be placed near to C_2, the centre of the eye-piece, and assumed that the image seen through

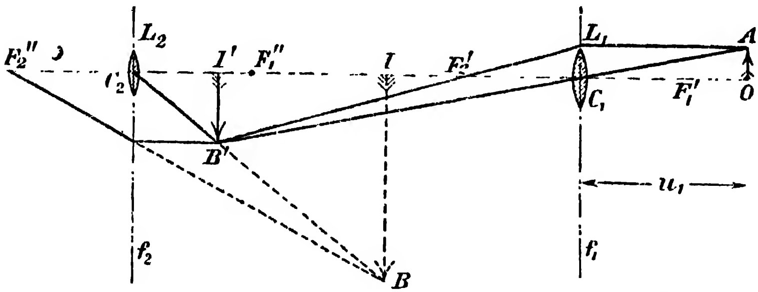


FIG. 25·11 (*a*).—Principle of an Astronomical Telescope *not* in Normal Adjustment.

the instrument is at the nearest distance of distinct vision, an expression for the angular magnification identical with that given on p. 511 for a compound microscope is obtained. The

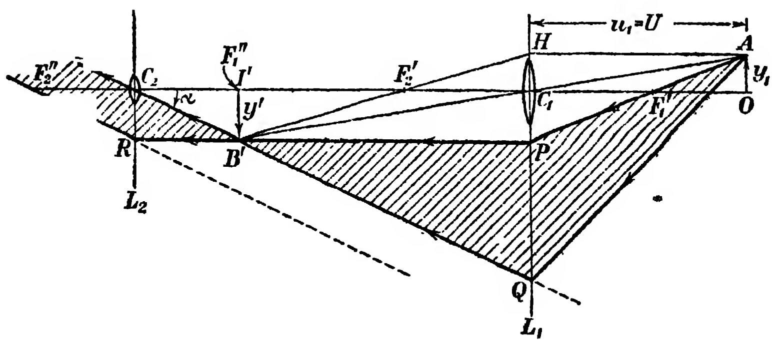


FIG. 25·11 (*b*).

systems differ only in the values of the focal lengths of the lenses and the relative position of the object to the objective. The expression so obtained is not a convenient one. It can be made much more convenient if it is stipulated that the eye shall be assumed to be at the objective when the usual angle for the object

is calculated. In addition, let the image seen through the instrument be situated at infinity. Then, cf. Fig. 25·11 (*b*),

$$M_{v\to\infty} = \frac{|I'B'|}{F_2} \div \frac{|y_1|}{|C_1O|} = \frac{U}{F_2}\cdot\frac{|y'|}{|y_1|}$$
$$= \frac{U}{F_2}\cdot\frac{|C_1I'|}{U} = \frac{|C_1I'|}{F_2}$$
$$= \frac{U}{U - F_1}\cdot\frac{F_1}{F_2},$$

when the value for C_1I' is expressed in terms of U and F_1. When the object is at infinity, i.e. the telescope is in normal adjustment, the magnifying power becomes equal to $\frac{F_1}{F_2}$.

Experimental Determination of the Magnifying Power or Angular Magnification of a Telescope. ***Method i:*** Set up at one end of the laboratory a long piece of inch graph paper on which the inch lines have been heavily marked in ink. At the other end of the room set up a converging lens of long focal length to form a real image of the scale: the position of this image should be located by means of a pin so placed that there is no parallax between it and one of the divisions in the image of the scale. Then arrange a second converging lens of shorter focal length so that it produces a magnified virtual image of the first image. This should be viewed with ***both*** eyes open and the eye-lens adjusted until the scale and its image can be seen at once. By adopting this method one is quite certain that the plane of the image coincides with that of the object itself. The angular magnification of the system is then equal to the ratio of a certain number N of divisions as seen directly which appear to coincide with *n* of the divisions as seen through the telescope.

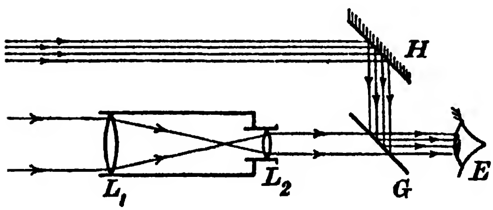


FIG. 25·12—Apparatus for Determining the Magnifying Power of a Telescope.

The difficulty often experienced of viewing the scale directly with one eye and through the telescope with the other at the same time may be overcome with the aid of the apparatus shown in Fig. 25·12. G is a thin sheet of glass arranged at 45° to the axis of the system; it permits the rays coming through the telescope to enter the eye E. M is a plane mirror reflecting direct rays from the distant scale on to G; a portion of this light is

reflected into E and the two images are seen superposed. A value for the magnification is then easily obtained.

Method ii: Focus the telescope for parallel rays and then direct it towards a white cloud. Receive the emergent light—which is a parallel beam—on a piece of ground glass placed at right angles to the axis of the telescope. A bright circular patch of light known as the '***Ramsden Circle***' will appear on the glass. Measure the diameter of this circle—D_2—and the diameter of the objective—D_1. Then $M = D_1 \div D_2$ [proved below].

Method iii: Place a diaphragm over the objective, the diaphragm being pierced with two small holes at distance r_1 apart. Receive the images of these on a piece of ground glass and measure their distance apart—r_2—by means of a travelling microscope. Then $M = r_1 \div r_2$.

Method iv: Focus the telescope for parallel rays and direct it towards the sky. The distance between the two lenses is then $|f_1| + |f_2|$, where f_1 and f_2 are the focal lengths of the objective and eye-piece respectively. Then remove the objective and obtain on a suitably placed screen the image of the aperture produced by the eye-piece. Let d_2 be its diameter. If d_1 is the diameter of the aperture, $M = d_1 \div d_2$.

Proof: Since the object is at a distance $|f_1| + |f_2|$ from the eye-piece, the image distance, v, is given by

$$-\frac{1}{|v|} - \frac{1}{|f| + |f_2|} = -\frac{1}{|f_2}$$

$$\therefore |v| = \frac{|f_2|[|f_1| + |f_2|]}{|f_1|}$$

But $$\frac{d_1}{d_2} = \frac{|u|}{|v|} = \frac{|f_1| + |f_2|}{|v|} = \left|\frac{f_1}{f_2}\right| = \text{M}.$$

The Terrestrial Telescope.—In viewing terrestrial objects through an astronomical telescope inconvenience is often caused by the fact that the image is inverted. To overcome this difficulty the terrestrial telescope shown in Fig. 25·13 may be used. We shall assume that the telescope is in normal adjustment, i.e. the object and final image are both at infinity. Parallel rays incident on the lens L_1 are brought to a focus in the focal plane of this lens so that if we imagine that the object extends from a point on the axis of the system to a point P from which the parallel rays considered emerge, I_1B_1 will be the real image of the object produced by L_1. This image is inverted, but it is made erect by means of a converging lens L_2. This is arranged in such a position that the distance from it to I_1 is twice its focal length. The real image which it produces is I_2B_2 and although no additional magnification has been achieved

by the arrangement adopted, the image is now erect and the distance I_1B_2 is a minimum consistent with the focal length of L_2. This latter condition is advantageous since the total length of the system cannot be increased beyond definite limits without causing the

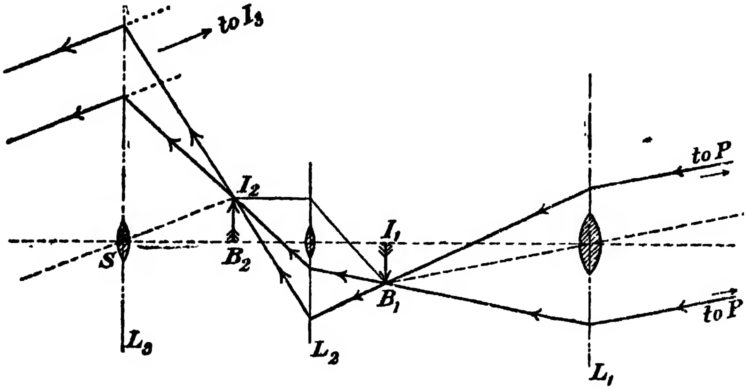


FIG. 25·13.—Principle of a Terrestrial Telescope in Normal Adjustment.

system to become unwieldy. A third lens L_3 is placed so that I_2B_2 is in its first focal plane: the final image is then at infinity, the direction of the emergent beam being parallel to the secondary axis

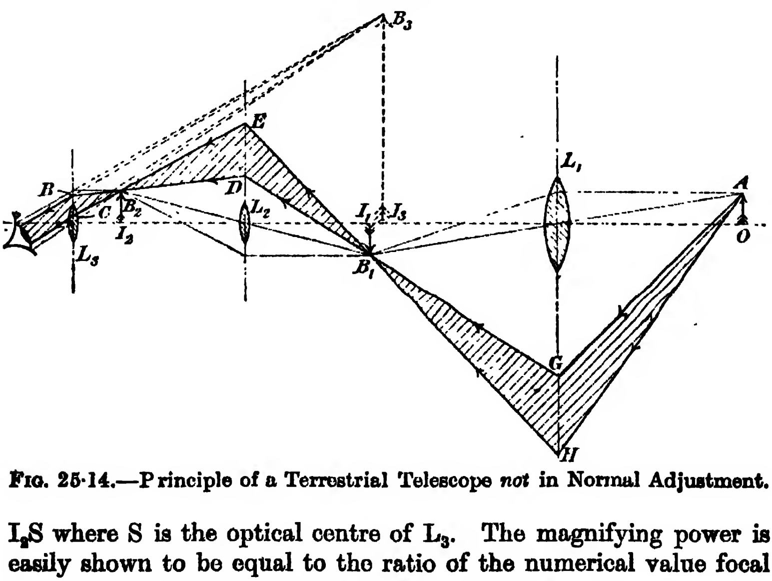


FIG. 25·14.—Principle of a Terrestrial Telescope *not* in Normal Adjustment.

I_2S where S is the optical centre of L_3. The magnifying power is easily shown to be equal to the ratio of the numerical value focal length of the objective to that of the eye-piece.

For a terrestrial telescope not in normal adjustment Fig. 25·14

indicates the method of locating the final image. To construct the path of the rays through the system by means of which an eye sees some particular point in the object, the corresponding point in the image is joined to the periphery of the pupil. Let B and C be the points at which these lines intersect the central plane of the lens L_3. Those portions of the lines from B and C to the eye are shown in full since they represent actual rays, while those portions drawn to the point in the image are dotted since the image is virtual. The points B and C are then joined to B_2 and produced to intersect the principal plane of L_2 in D and E. These points are joined to B_1 and DB_1 and EB_1 produced to meet the principal plane of the lens L_1 in G and H. By joining these two points to A we have traced the rays from A through the system to the eye.

Galileo's Telescope.—The disadvantage of the astronomical telescope when used to view terrestrial objects has been overcome as described above by the use of a third lens. The objection to this is that the length of the telescope has been increased. Galileo's telescope has the advantage that it produces an erect image and yet the distance between the lenses is less than in an astronomical telescope having an equal objective and magnifying power. Let us consider Galileo's telescope when in normal adjustment as shown in Fig. 25·15. Rays proceeding from a point in the object in a direction

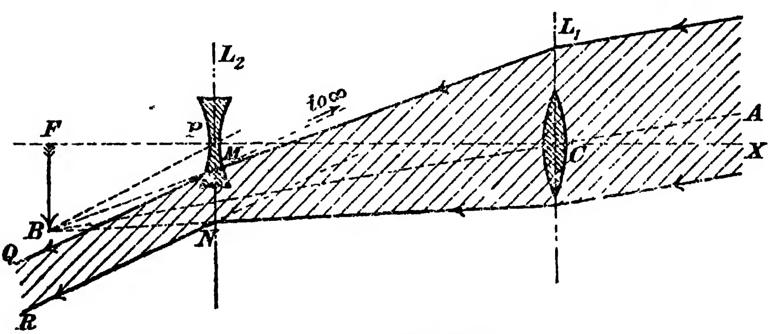


FIG. 25·15.—Principle of Galileo's Telescope in Normal Adjustment.

parallel to AC, a secondary axis of L_1 would, in the absence of the eye-piece, be brought to a focus at B, that point in the second focal plane of L_1 where it is intersected by AC produced. The eye-piece, L_2, is a diverging lens of short focal length so placed that its first principal focus is also at F. When the converging beam of rays from L_1 is refracted by L_2 the emergent rays are parallel to the secondary axis PB of the eye-piece [cf. p. 428]. The final image is therefore a virtual one situated at infinity: it is erect. The angular magnification, M, of this telescope, which is the ratio of

the angle subtended at the eye by the image to that subtended by the object, is given by

$$M = \frac{\widehat{BPF}}{\widehat{ACX}} = \frac{\widehat{BPF}}{\widehat{FCB}} = \frac{|f_1|}{|f_2|}.$$

The angular magnification for a given Galilean telescope is determined experimentally as for an astronomical telescope [cf. p. 516, method (*i*)].

Fig. 25·16 shows how the image is produced when Galileo's telescope is *not* in normal adjustment.

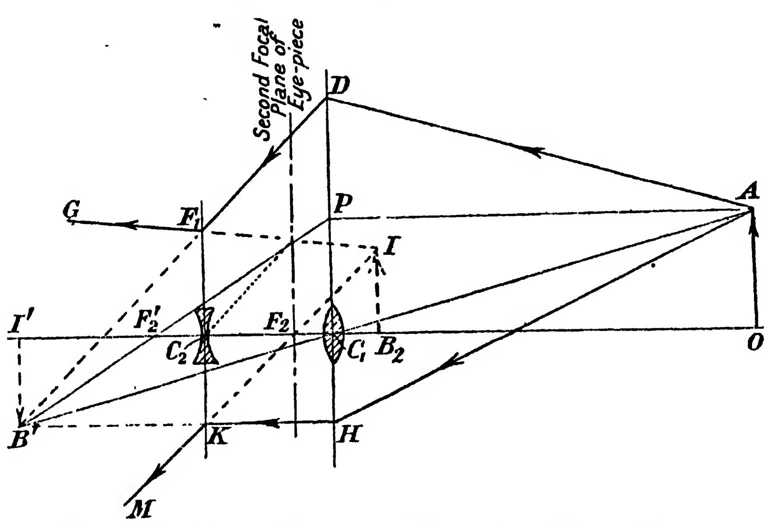


FIG. 25·16.—Principle of Galileo's Telescope *not* in Normal Adjustment.

Prism Binoculars.—The field of view in Galileo's telescope is not uniformly bright and for a magnifying power 3 [the usual value when the telescope is made in the form of opera-glasses] the field of view is 2·5 times smaller than that of an astronomical telescope having the same power. The reason for this deficiency in a Galilean telescope lies in the fact that for such an instrument the Ramsden circle is virtual. When, as for an astronomical telescope, the Ramsden circle is real, the eye may be placed at that position so that oblique pencils from the edges of a wide field are received by it. With a virtual Ramsden circle the eye cannot be made to coincide with it and consequently the pencils of light from different parts of the field are divergent as they leave the eye-piece and only a few of them enter the eye of an observer. To overcome these defects in a Galilean telescope prism binoculars have been designed. The essential difference between such an instrument and an astro-

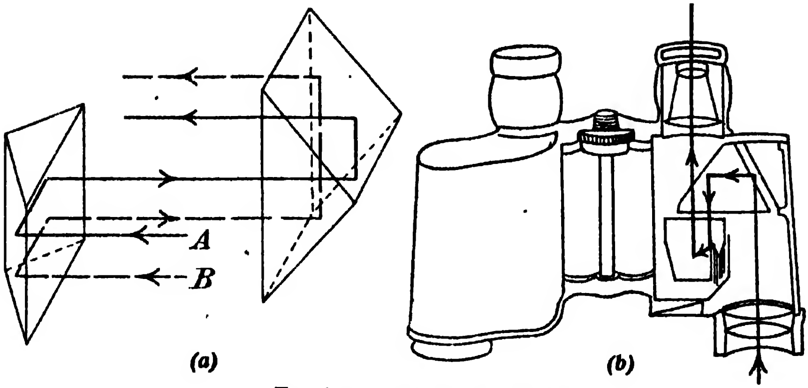


FIG. 25·17.—Prism Binoculars.

nomical telescope is that two right-angled glass prisms with their edges at right angles to each other are placed between the lenses. If two parallel rays A and B, Fig. 25·17 (*a*), strike the base of the first prism, the refracting edge of which is vertical, the rays enter the prism and are reflected from one face to the other and then again, so that they finally emerge parallel to their original direction but with lateral inversion, i.e. the right-hand side is now the left, and vice versa. If these rays fall on a second prism whose refracting edge is horizontal the two emerging rays are inverted as shown. Fig. 25·17 (*b*) illustrates a modern form of field glass or prism binoculars.

Newton's Reflecting Telescope.—When Newton discovered that the images produced by lenses were always indistinct at their edges he ceased to try to improve Galileo's invention and designed an instrument in which the refraction of light was avoided. The principle underlying this design is indicated in Fig. 25·18. A concave mirror, of focal

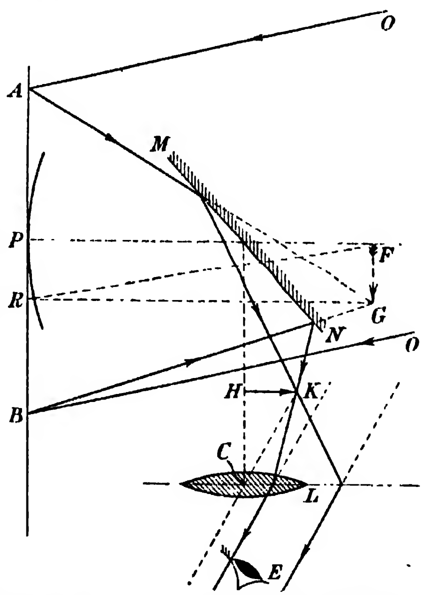


FIG. 25·18.—Principle of Newton's Reflecting Telescope in Normal Adjustment.

length $f_1 = F_1$ (say), acts as the objective and reflects parallel rays OA, OB, from the uppermost point in a distant object to a point G in the focal plane of the mirror. The point G is determined by constructing FR parallel to the incident rays. Since the ray FR passes through the focus of the mirror it will travel along RG after reflexion, where RG is parallel to the axis of the mirror. If the object has its lowest point on the axis, then GF will represent the image produced by the mirror alone. To avoid the necessity of looking at this image directly and thereby obstructing some of the incident rays, Newton placed a plane mirror, MN, at 45° to PF, so that a brilliant image was formed at KH. To find the position of this image we note that K and H are at the same perpendicular distances from the mirror as are the points G and F respectively. If this image lies in the first focal plane of a converging lens L, of focal length $f_2 = -F_2$ (say), the final image is at infinity. The telescope is then in normal adjustment. An expression for the magnifying power, M, may be obtained as follows.

$$M = \frac{\text{angular size of image}}{\text{angular size of object}}$$

$$= \frac{|HK|}{F_2} \div \frac{|GF|}{F_1}$$

$$= \frac{F_1}{F_2} \qquad [\because HK = FG]$$

Some Modern Reflecting Telescopes.—Since it is difficult to produce large lenses it seems likely that future improvements of telescopes must be made with those of the reflector type. In viewing faint stars, for example, it is necessary to have as large an aperture as possible since more light then enters the telescope and so produces a brighter image. The largest objective at present in use is that at the observatory at Lake Geneva, Wisconsin, U.S.A.: its diameter is 40 inches. The largest concave mirror is the objective of the 100-inch Hooker telescope at Mount Wilson, California, and any imperfections which may exist in the interior of its glass and which would be fatal if that glass constituted a lens, become of minor importance when the light is merely reflected from the surface of the glass.

Some attempts have been made to construct telescopes of this type in which the concave mirror consists of a pool of mercury rotating uniformly about a vertical axis through its centre. The surface of the mercury assumes a parabolic form under such circumstances so that a point image of a distant source is obtained in the focal plane of the mirror.

The Erecting Prism.—Let us suppose that a lens L, Fig. 25·19, produces a real inverted image, I_1B_1, of an object OA. A real erect image may be obtained in the following way. PQR is a glass prism in which the angle at P is a right angle. Let us consider the extreme rays AC and AK of the cone of rays proceeding from A and which pass through the lens. In the absence of the prism the rays AC and AK are refracted along the paths CB_1 and KB_1. When the prism is placed in position, the ray CD is refracted along the path DE and is then reflected along EF. Finally this ray emerges along FB. The ray KM is similarly refracted and reflected along the path

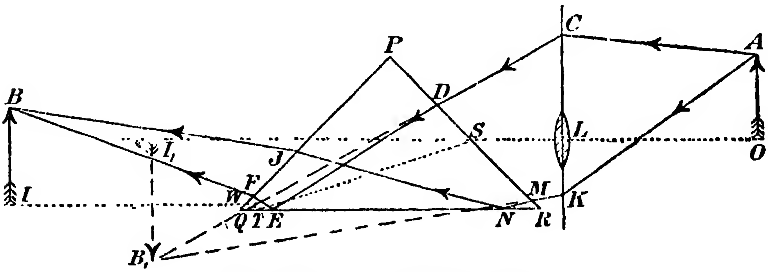


FIG. 25·19. An Erecting Prism.

KMNJB. A real image of the point A is now formed at B. Similarly, the ray OLS pursues the path STW through the prism and emerges as WI, so that IB is the image produced by the lens and prism together. It will be noticed that I_1B_1 and IB are not at the same distance from L [Verify by performing an actual experiment.]. Moreover, O, I_1 and I are not necessarily collinear.

If OA is a lantern slide, inserted the correct-way-round for normal showing, then IB will be an erect image but with lateral inversion —it is only when the slide is erect and inserted the wrong-way-round that IB is erect and without lateral inversion.

The Periscope.—Suppose two plane mirrors, M and N, Fig. 25·20, are arranged so that rays of light incident upon M are reflected so that they fall upon N from whence they are reflected in a direction parallel to the incident rays. For this to be possible the mirrors must be parallel. A glance at the diagram shows that the rays have suffered a lateral displacement. It is therefore possible for an observer to see objects by looking into the mirror N without himself being seen. This is the essential principle of a periscope, only the range of vision is increased by combining it with a telescope. For simplicity we will assume that an astronomical telescope is used. The two plane mirrors do not invert the image, yet when they are combined with such a telescope the final image will be inverted since this is a characteristic feature of an astronomical telescope. Some piece of additional apparatus must therefore be inserted in the system. Let us suppose that an erecting prism has been placed in front of M as in Fig. 25·21. Parallel rays from a point A in a distant object pass through the prism

and strike the mirror M at B and C whence they are reflected along BD and CE. If OH is the secondary axis of the lens L_1 parallel to BD and CE these rays will be brought to a focus at H, the point in the second focal plane of the lens where it is cut by OH. An image is therefore produced at HI_1. A virtual image of this is produced by reflexion in the mirror N, and if this image is at KI_2 in the first focal plane of the lens L_2 the final image is at infinity. To complete our trace of the rays from A through the system, the point K, the image

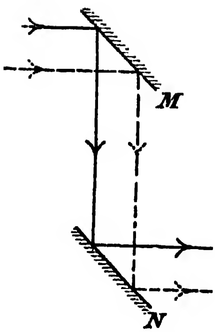


FIG. 25·20.

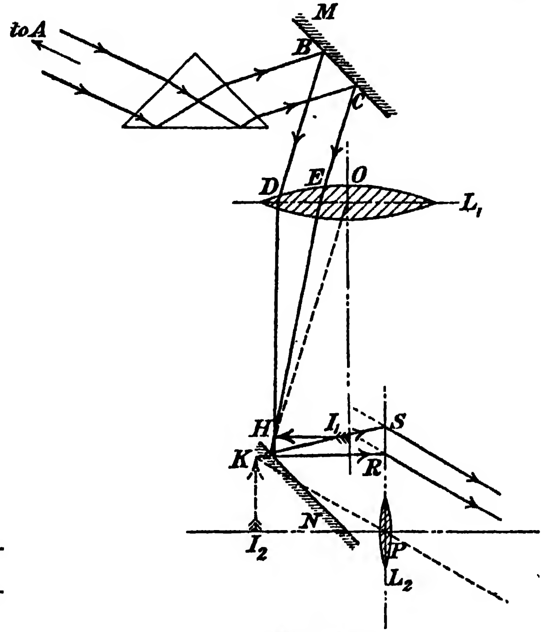


FIG. 25·21.

Principle of the Periscope.

of H, is joined to the points where DH and EH meet N, and the lines are produced to cut the principal plane of the lens in R and S. The rays are then refracted by the eye-piece L_2 so that they proceed in directions parallel to KP, the secondary axis of L_2 passing through K.

In practice the prism is not placed in front of the plane mirror M, but it has been drawn in that position since if the student will carry out the above construction it furnishes an excellent exercise in the principles of geometrical optics. In actual periscopes the erecting device is placed after the rays have passed through the lens L_1.

Diascope or Optical Lantern.—The essential features of a diascope, an optical instrument used for projecting images on a screen, are indicated in Fig. 25·22. A 'Pointolite' lamp, S, is placed at a short distance from a large converging lens L_1, termed a condenser. Sometimes a water trough is placed before this lens to reduce the amount of heat radiation upon it and so render it

less liable to fracture. In the absence of a condenser the amount of light incident upon the slide is confined to the cone ASB, whereas when the condenser is used the light in the cone CSD illuminates the slide if it is suitably placed. An achromatic lens L_2 is erected

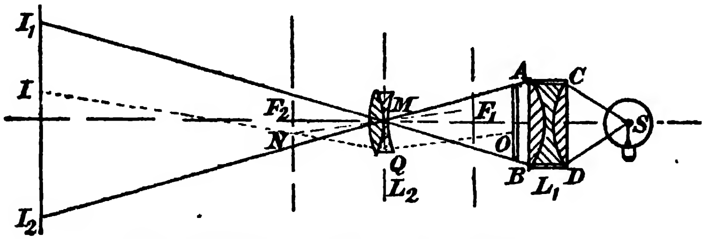


FIG. 25·22.—Principle of a Diascope or Optical Lantern.

in front of AB and their distance apart varied until a clear image I_1I_2 is obtained. The path of a ray OQ proceeding from a point O in AB is constructed in the usual way by drawing the secondary axis MN parallel to OQ. Then QNI is the path of the ray after leaving the lens.

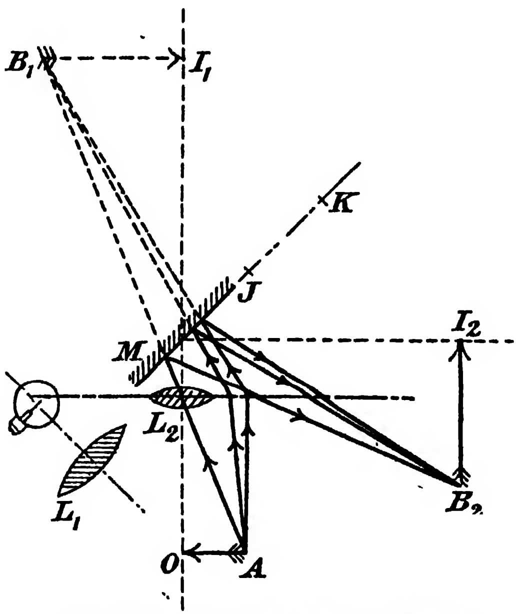


FIG. 25·23.—Principle of an Epidiascope.

The Epidiascope.—An epidiascope is an arrangement of lenses and a mirror for projecting an image of an opaque object on a screen. A converging lens, L_1, Fig. 25·23, is placed so that its second focal

plane contains the filament of an electric bulb. This filament lies in a plane perpendicular to that of the paper so that a maximum amount of light may pass through L_1. This light illuminates any object such as OA. L_2 is the projecting lens carried in a suitable stand to enable its distance from the object to be varied. In the absence of the plane mirror M which consists of a piece of optically worked plate glass silvered on its front surface to avoid the formation of multiple images [cf. p. 415], a real image I_1B_1 would be formed. When the mirror is in position the final image is at I_2B_2 which is located as follows: from B_1 and I_1 erect perpendiculars to the plane of the mirror M and produce them to B_2 and I_2 respectively such that $B_1J = JB_2$, etc. Then I_2B_2 is the image, and the path of the rays from A to B_2 is completed in the usual way.

The Telemeter or Range Finder.—It is at once apparent from trigonometrical considerations that if the base of a right-angled triangle is known as well as the angle which the base subtends at an object placed at the apex of the triangle, then the distance of the object is readily calculated. If the distance is very great and the base small and normal to the line of sight the distance is again calculable from the length of the base and the circular measure of the angle subtended. To discover the magnitude of this angle many forms of telemeter or range finder have been invented: the principle of the Barr and Stroud range finder is illustrated in Fig. 25·24. Rays of light PQ and RS inclined to each other at an angle θ are incident upon two mirrors

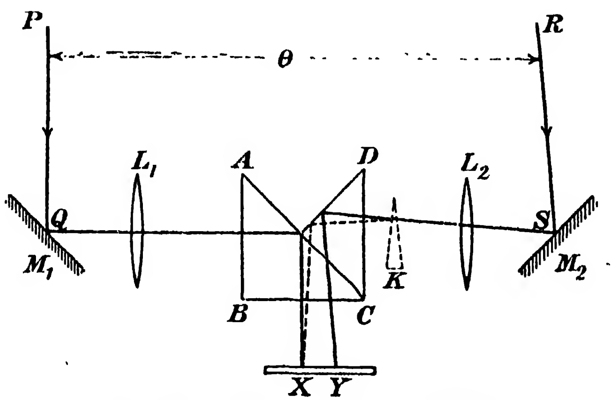


Fig. 25·24.—Telemeter or Range Finder.

M_1 and M_2 inclined at 45° to the 'base.' Converging lenses L_1 and L_2 then produce images of the object at their respective second foci and the specially constructed prism ABCD enables these images to be seen at X and Y when the small prism K is absent. In the sketch X and Y are represented as points on a screen: actually they are formed in the first focal plane of a microscope. Now the distance XY is a measure of θ, the angle required. Instead of measuring the distance XY the small prism K is moved until the two images coincide. The path of the central ray from M_2 after leaving K is shown by the dotted line. A pointer attached to K and moving over a scale parallel to the base of the instrument gives the distance of the object directly. The scale is calibrated by sighting objects at known distances.

Telescope and Microscope Objectives.—In our treatment of optical instruments we have always supposed them to be fitted with single lenses, i.e. the objective and eye-piece are each a simple lens. Such lenses suffer very considerably from defects known as chromatic aberration and spherical aberration. The objective of a refracting telescope is corrected for chromatic aberration by combining a converging lens of crown glass with a diverging lens of flint glass, but so that the combination still acts as a converging lens [cf. Fig. 25·25]. To reduce spherical aberration in it the lens is mounted with its converging component towards the object. But the use of two lenses has brought with it a disadvantage which is overcome in the following way:—If the inside faces of the two lenses are separated from each other some of the light passing through the converging lens will be reflected from the front face of the second lens with a consequent reduction in the brightness of the image. This defect is eliminated by making the radii of curvature of the inner faces of the two lenses identical, and cementing them together with Canada balsam, the refractive index of which is intermediate between those of crown glass and flint glass.

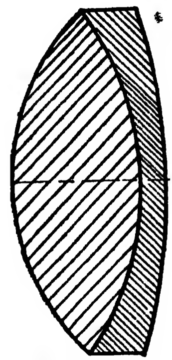


FIG. 25·25.—Telescope Objective.

The field lens of a telescope is made large so that the amount of light collected by it shall be as large as possible in order that the final image shall be bright; also, that details in the object shall be clearly seen—we say that the ***resolving power*** of the instrument has been increased. Moreover, it has been shown that the magnifying power due to a telescope, in normal adjustment, for example, is expressed by the ratio

$$\frac{\text{focal length of objective}}{\text{focal length of eye-piece}}.$$

It would therefore appear that by increasing the relative focal length of the objective the magnifying power could be increased indefinitely. Now although the magnifying power may be increased in this way, no advantage is gained for no further details become visible, unless the diameter of the lens is increased: for a high resolving power requires a lens of large diameter.

The objective of a first-class microscope is a very complicated piece of apparatus; it is difficult to construct and therefore expensive. For very high-power work it is necessary to immerse the specimen in cedar-wood oil, the front surface of the objective also being in the oil. Such a lens is known as an ***immersion***

lens. A diagrammatic representation of Abbé's immersion objective is shown in Fig. 25·26. The lowest lens L_1 is a plano-hemispherical convex lens; the intervening space between this and the object is filled with the oil, which has the same refractive index as glass. This implies that no refraction takes place until the rays leave L_1. If O, a point in the object, is such that O and O_1 are ***aplanatic points,*** i.e. O_1 is a point image of the point source O, with respect to A, the point where the hemispherical surface cuts the principal axis of the lens, then all rays from O appear to diverge from O_1, irrespective of the obliquity of the rays. A second lens L_2 is placed above L_1, its lower face B having O_1 as its centre of curvature. The rays which apparently proceed from O_1 are not refracted at B, but only at the upper face C of the lens L_2, so that they appear to proceed from O_2, a point which is made aplanatic with respect to O_1. Compound lenses L_3 diminish the effects of chromatic aberration, i.e. they are composed of convex lenses of crown glass, cemented to concave lenses of flint glass.

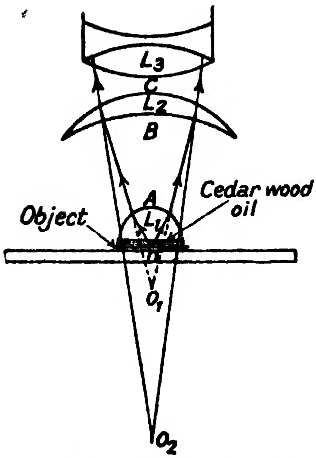


FIG. 25·26.—Abbé's Immersion Objective.

Aplanatic Foci.—Let O and I, Fig. 25·27, be two points on the axis AC produced of a convex spherical refracting surface such that $CO = R/\mu$ and $CI = \mu R$, where μ is the index of refraction of the material and R the radius of curvature of the surface. The Δ's COP and CPI are similar, since

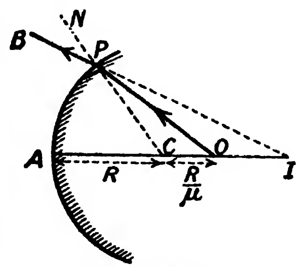


FIG. 25·27.

$$\frac{OC}{CP} = \frac{1}{\mu} = \frac{CP}{CI} \text{ and } \widehat{OCP} = \widehat{PCI}.$$

Hence
$$\frac{\sin CPI}{\sin CPO} = \frac{\sin COP}{\sin CPO} = R \div (R/\mu) = \mu.$$

It therefore follows that IP is the direction of the refracted ray and hence, whatever ray proceeds from O, the direction of the refracted ray passes through I, i.e. O and I are ***aplanatic foci.***

EXAMPLES XXV

1.—How may a converging lens and a diverging lens be employed to form a telescope? Give a carefully drawn diagram of the paths of the rays by means of which an eye may observe a point in a distant object.

2.—State the arrangement of lenses necessary for the formation of an opera glass. Show by means of a diagram how the position of the final image may be found. Trace the path through the system of rays from a point on a distant object off the axis of the system.

3.—The picture on a lantern slide 3 in. square is to be projected upon a screen 18 ft. distant from the slide by means of a lens of 10 in. focal length. At what distance from the lens must the slide be placed, and what will be the size of the picture on the screen?

4.—Describe the optical parts of a compound microscope and trace the rays through the system by means of which an eye sees a point in an object off the axis of the microscope. Upon what does the magnifying power of a microscope depend? How would you measure it?

5.—Describe the action of a compound microscope formed by two convex lenses and show with an example how to determine its magnifying power. Will this last be affected by short-sightedness in the observer. (L. '25.)

6.—What is an achromatic lens? Give an account of the principles of construction of achromatic prisms and lenses. (L. '24.)

7.—The focal lengths of the lenses of a reading telescope are 25 cm. and 4 cm. and it is used to view a scale 1 metre from the object-glass. If the image is formed 25 cm. from the eye, which is close to the eye-piece, draw a diagram showing the paths of the rays through the telescope. Calculate the magnifying power of the instrument. (L. '30.)

8.—Give a general explanation of the construction of an achromatic lens suitable for use (*a*) as a telescope objective, (*b*) in a photographic camera. How does the appearance of the image seen through an astronomical telescope vary with the diameter of the objective?

9.—Explain how you would arrange three converging lenses on a common axis so that a beam of light from an object on the axis and outside the system will produce (*a*) an erect real image, (*b*) an erect virtual image, after passing through the three lenses. Give diagrams showing the paths of rays of light from a point on the object to the corresponding point on the image in each case.

10.—A model of a compound microscope is made up of two thin converging lenses of focal lengths 3 cm. and 9 cm. respectively, the interval between the lenses being 25 cm. Where must the object be placed so that the final image may be at infinity? What will be the magnifying power if the microscope, as thus arranged, is used by a person whose least distance of distinct vision is 30 cm.?

CHAPTER XXVI

PHOTOGRAPHY

The Pin-hole Camera.—The simplest form of camera—a device for producing an image of an object—consists of a very small hole pierced in a light-tight wooden box. If a piece of white paper is placed at a short distance away from the hole an image of the external object is produced on it. The formation of an image under these conditions is explained on the hypothesis that light travels in straight lines. If a gas-filled electric lamp with its irregularly-shaped tungsten filament, Fig. 26·1, is placed in front of a piece of tin-foil in which a small pin-hole has been pierced, light travels in straight lines from the different points of the filament,

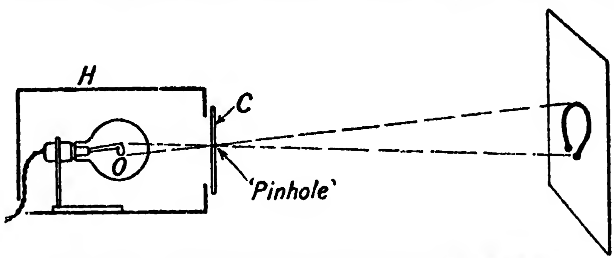


FIG. 26 1.—The Principle of a Pin-Hole Camera.

passes through the hole, and falling on the screen behind, produces a series of point images—the complete picture is an inverted image of the filament. [The lamp should be screened by a suitable box.] This image becomes more and more blurred as the size of the hole is increased, because the light which passes through each portion of the hole gives rise to an image. A blurred image is also obtained if several holes are pricked in the foil; each hole gives rise to an image and the blurred nature of the picture is caused by the overlapping of the several images. It must be noted, however, that the shape of the hole does not affect the image, providing, of course, that the hole is small.

The Photographic Camera.—The great objection to the use of a pin-hole camera, if sharp images are to be obtained, lies in the fact that only a small amount of light is available for the production

of the image—because the pin-hole is small. On the other hand, the dimensions of the image are strictly proportional to those of the object. If the pin-hole is replaced by a single bi-convex lens, then the amount of light available is increased many times. But the insertion of a lens has introduced all the errors to which such a lens is subject—spherical aberration, chromatic aberration, distortion, astigmatism, lack of flatness of field, etc. Hence, in order to produce a picture which shall be more true to real life, the optician has designed a lens in which these defects are reduced to a minimum; they can never be reduced to zero, although the modern anastigmatic lens is an excellent example of the optician's skill.

The Photographic Objective.—In designing this lens so that it shall be as free as possible from chromatic aberration it has to be remembered that the conditions under which it is to be used are very different from those under which a telescope objective is employed. In the first place, the eye is used to decide whether or not the image on the focusing-screen is sharp. Since the eye is most sensitive to yellow light, whereas the chemicals in the emulsion of the film are generally most sensitive to the blue and actinic rays, it follows that the focal lengths of the lens for the yellow and blue rays should be identical, for although the yellow rays are less actinic than the blue, the yellow and blue rays must be focused in the same plane, for, otherwise, although the image as seen by the eye may be judged sharp, that obtained on the plate will be fuzzy since other rays have been more responsible for its production.

The Photographic Plate.—In the chapter on dispersion it has been stated that ordinary white light is composed of several colours which are capable of being separated out into a spectrum by means of a prism. It has long been known that light of any colour is capable of producing a chemical change—the light is said to act photochemically. To such an action must be attributed the tanning of the skin after prolonged exposure to the sun, and the change in the colour which occurs when pigments are similarly exposed.

The darkening in colour of silver salts under the influence of light is a fact which has been well established for many years; the early discoverers of this phenomenon were puzzled by the appearance of something dark which had to be attributed to light. During the process of blackening silver chloride some free chlorine is evolved, for it has been shown that chlorine water, if applied to some darkened silver chloride, restores the original colour. Many writers have maintained that a subchloride of silver is produced which combines with the free silver chloride to form a complex compound, $AgCl . xAg_2Cl$. In the manufacture of modern photographic plates silver bromide is used, and the influence of light upon it is similar to the action upon the chloride.

To make a photographic plate or film a piece of glass, or sheet of celluloid, is coated with a film of gelatine carrying particles of silver bromide in suspension. When dry it is ready for use.

The Latent Image, and its Development. If such a plate is exposed to light rays, e.g. the image of some illuminated object is allowed to fall upon it, the bright portions of the image cause a greater blackening of the bromide than do the darker portions of the image. The effect produced on such a plate is not visible; there is only present the ***latent image,*** and a developer is used in order to render this image visible. The developer consists of a reducing agent, such as ferrous sulphate or pyrogallic acid, which converts the bromide particles which have been affected by the light into metallic silver, which is deposited in the form of black granules. When the black granules have become sufficiently dense, the plate is removed from the developer, washed, and placed in a solution of sodium thiosulphate or hypo ($Na_2S_2O_3 . 10H_2O$). The function of this salt is to dissolve the unaffected portions of the silver bromide remaining on the plate, so that the ultimate result is a distribution of black metallic silver particles throughout a gelatine film. The distribution varies according to the manner in which the light and shade were distributed in the original subject. This final record is called a ***negative,*** and the negative is perfect when the contrasts in the subject have been recorded faithfully. To prepare a true likeness from such a negative, a ***positive*** must be made. Paper, treated similarly to the original plate, is placed in immediate contact behind the negative [enlargers being omitted] and the whole exposed to a uniform light. The light traverses the transparent portions of the film more readily than elsewhere, so that a developable image is produced on the sensitized paper. After development and fixing, a permanent photograph is obtained.

Orthochromatic Plates and Films. Light Filters. If a blue and a red object are observed together by a normal eye the blue one may appear to be darker than the red one; in a photograph, however, the red will appear to be darker if an ordinary plate is used. The reason for this lies in the fact that an ordinary photographic emulsion is more sensitive to blue than to red light; were it not so, a ruby lamp could not be used in the dark room. Similarly, if a landscape is photographed, the beauty of the original does not survive in the negative, for the plate fails to differentiate between the varying shades of green, while the clouds may not be retained at all; the dark blue of the sky cannot be distinguished from the white clouds because the blue and white rays are equally actinic.

In an orthochromatic or isochromatic plate or film the emulsion is made sensitive to yellow and green rays, but at the same time it remains exceptionally sensitive to the blue and violet rays. If,

therefore, the aim of the photographer is to obtain a true monochrome picture of the object he must cover his lens with a filter. A filter is a piece of stained gelatine (yellow) fitted between two pieces of glass, or, better still, the glass itself is stained. Now the function of this yellow filter is to absorb some of the blue and violet rays. If the grade of filter has been properly selected then the transmitted rays will be such that the resulting negative can yield a picture which shall be almost as pleasing as the original object.

Some makers of orthochromatic plates place a yellow dye in the emulsion of the plate when the use of a filter is not necessary. Such self-screening plates are very effective, but no isochromatic plate will give such good results as a panchromatic plate.

Panchromatic Plates. The isochromatic plates mentioned above are still insensitive to the red rays which emanate from the object, so that the red portions of the object assume, in the print, a tone which is much too dark. Now panchromatic plates are sensitive to all colours, so that it is impossible to develop them with the aid of a red lamp. Like orthochromatic plates, however, they are still exceptionally sensitive to the blue region of the spectrum. If, therefore, the full benefit is to be obtained from such plates a yellow filter must be used. The darker the filter the longer the exposure, but the resulting negative is better provided that the plate has not been unduly over-exposed.

It may be thought that the development of a panchromatic plate is very difficult since it must be done in darkness. Fortunately this is not so, especially if the plate is first desensitized. The plate is removed from the camera slide in complete darkness and then placed in a dilute solution of pinacryptol green for one minute. After this the plate may be developed in a yellow light or by the aid of a fairly distant candle flame.

Contrast Photography—Clouds.—Whilst panchromatic plates have been designed to render correctly the tones present in an object, they can also be used to accentuate certain details in it. If the colours red and blue predominate, then a correct rendering is obtained by a filter of such colour that about four times the normal exposure is required. If, however, for any reason it is necessary to contrast the red and blue then a red filter is placed over the lens. This red absorbs the blue rays entirely so that the corresponding parts of the negative are not affected. In this way the contrast is accentuated.

CHAPTER XXVII

THE VELOCITY OF LIGHT

Astronomical Method.—All the attempts made by Galileo and others having failed to fix a definite value for the velocity of light, it was assumed that the speed of light was infinite until some curious results were obtained in 1676 by a Danish astronomer, RÖMER, with reference to the periodic times of the satellites or moons revolving round the planet Jupiter. These could only be explained by assuming that the velocity of light was finite. From observations on the times of successive disappearances of the innermost satellite—the one which moves in the same plane as that containing the sun, earth, and Jupiter—Römer predicted the times when future eclipses should occur. He discovered that his calculated times did not agree with those at which the eclipse actually took place and noticed that the discrepancy increased continuously as the earth moved away from Jupiter.

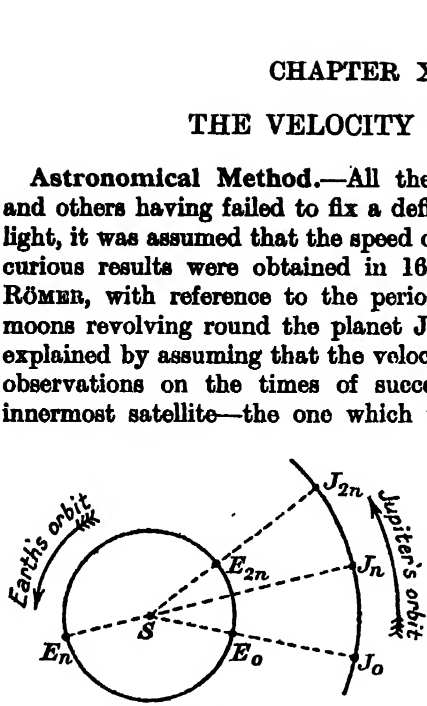


FIG. 27·1.—Römer's Method for Determining the Velocity of Light.

Let S, Fig. 27·1, be the sun while the positions of the earth and Jupiter at corresponding times as they each revolve round S are indicated by the letters E and J with appropriate suffixes. When the earth, sun, and Jupiter are collinear with the earth between the sun and Jupiter, the earth and Jupiter are said to be in conjunction: when they are collinear but the earth on the side of the sun away from Jupiter, they are in opposition. Let us suppose that an eclipse of Jupiter's innermost satellite takes place when the earth and Jupiter are at E_0 and J_0 respectively. If we measure time from the instant when the eclipse actually occurs, the disappearance will be seen at a time $\frac{E_0 J_0}{c}$, where c is the velocity of light across interplanetary space. Suppose that the next disappearance of

this satellite happens when the earth is at E_1 and Jupiter at J_1. If the period of revolution of this satellite is τ, the disappearance which actually takes place at time τ, will be observed at a time

$$\tau + \frac{E_1J_1}{c},$$

i.e. the true period cannot be directly observed. Let us further assume, however, that, when the earth and Jupiter are next in opposition at E_n and J_n, the satellite has made n revolutions. These will be complete at an actual time $n\tau$, but the completion will be observed on the earth at a time $n\tau + \frac{E_nJ_n}{c}$. The observed time, t_1, corresponding to n revolutions is therefore

$$t_1 = n\tau + \frac{E_nJ_n}{c} - \frac{E_0J_0}{c},$$

or $n\tau + \frac{2R}{c}$, if R is the radius of the earth's orbit.

Similarly n more revolutions will have occurred when the two planets are in conjunction at E_{2n} and J_{2n}, after the lapse of another 0·545 year. The observed interval between them will be

$$t_2 = n\tau - \frac{2R}{c}.$$

Hence
$$t_1 - t_2 = \frac{4R}{c}.$$

Römer found this difference to be 2,000 seconds, so that since $R = 93 \times 10^6$ miles, $c = 186{,}000$ miles per second, or 300×10^6 metres per second.

Fizeau's Method.—Fig. 27·2 is a diagram of the apparatus employed by Fizeau about 1849 to determine the velocity of light in air. An image of a powerful source of light, S, was produced by means of a converging lens L_1 and a glass plate, G, at F between two teeth of a toothed-wheel, W, rotating about a horizontal axis. In this wheel the teeth were made of precisely the same width as the interval between them. A converging lens L_2 was adjusted so that F was at its first principal focus, i.e. any light incident from F upon this lens emerged as a parallel beam. At a distance of 8·63 kilometres the receiving apparatus was erected. This consisted of a converging lens L_3 and an eye-piece L_4. The receiving apparatus was directed to pick up the light from the sending station so that an image was produced at P—the lens L_4 enabled this image to be observed and the collineation corrected. A plane mirror was then placed at P causing the light incident upon it to be reflected back to F. Some of this reflected light passed through the half-silvered mirror G to the eye-piece L_5.

When W was caused to rotate the light rays arriving at F from the source were alternately transmitted through a space between two teeth and then intercepted by a tooth. At slow speeds the light transmitted was able to travel from F to P and back before the wheel had moved even through a small angle so that an image was still seen. When the light was intercepted no image was seen. The effect of this slow rotation was to produce a succession of appearances and disappearances of the image, i.e. an eye at E perceived a flickering image provided that not more than 8 or 10 reappearances of the image occurred per second. When the speed of rotation was increased the flickering ceased owing to the persistence of images on the retina and the intensity of the image appeared to decrease continuously as the speed increased. Finally, a stage was reached when the field of view was dark—this meant

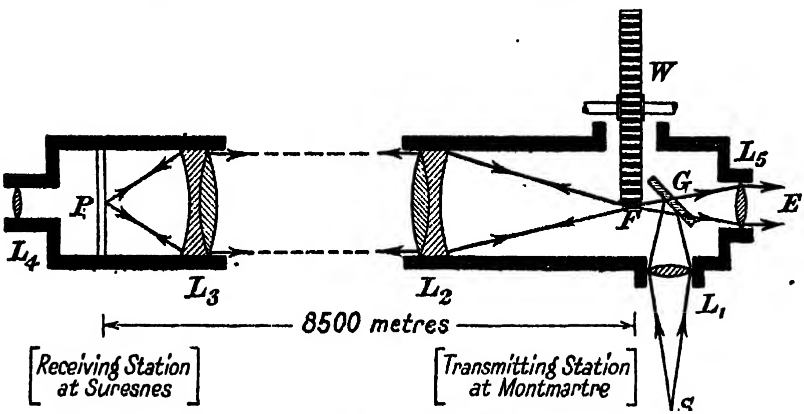


FIG. 27·2.—Fizeau's Apparatus for Determining the Velocity of Light.

that the time taken for light to travel from F to P and back was equal to that in which a tooth moved half the distance between consecutive teeth. The speed of the wheel having been measured by clockwork, the velocity of light was calculated as follows:—
Let l be the distance between the two stations, n the number of revolutions per second made by the wheel (12·6) when the image disappeared entirely, and N the number of teeth on the wheel (700). The time taken for the wheel to rotate so that each tooth moves into a position just occupied by the one in front of it, is $\frac{1}{Nn}$, i.e. $\frac{1}{2Nn}$ is the time in which light travels through air a distance $2l$. The velocity required is therefore

$$c = 2l \div \frac{1}{2Nn} = 4Nnl.$$

This method is open to the objection that it is difficult to decide exactly the instant when the darkness in the field of view is a maximum since the speed of the wheel could be varied between rather wide limits without producing any apparent change in the field. CORNU obviated this by using an electrical arrangement whereby the speed at any instant could be ascertained. He determined the speeds when the image first disappeared and also when it reappeared. A mean value was employed in the calculations.

Foucault's Method (1849–1862).—S, Fig. 27·3 (*a*) was a rectangular slit through which sunlight was passed. A fine wire was stretched across this aperture so that any image of it could be located with precision. This light was received by an achromatic converging lens, L, and the transmitted beam, after reflection at the plane mirror, M, was brought to a focus on a concave mirror R. The distance RM was equal to the radius of curvature of this mirror

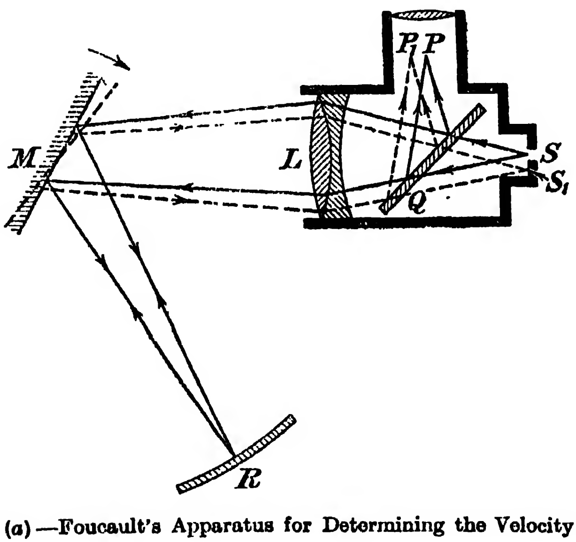


FIG. 27·3 (*a*)—Foucault's Apparatus for Determining the Velocity of Light.

so that the pencil of light incident upon it retraced its former path. This light was reflected from M and after passing through L formed an image coincident with the slit. Since it was impossible to observe the image under these conditions a parallel plate of glass was erected near S at an angle of 45° to the axis of the system. [This piece of glass was not half-silvered, as sometimes stated, for Foucault says 'La glace sans tain réfléchit une partie des rayons.'] A portion of the light reflected from M and incident upon this glass was reflected from it and produced an image at P. The

position of this image was observed by a micrometer eye-piece. The plane mirror M was capable of very rapid rotation about an axis perpendicular to the plane of the paper. Let us suppose that at some particular instant, when the mirror was rotating very slowly, it was in such a position for the light reflected from it to fall on R, from whence it retraced its path to M. Since the rotation of M has been assumed to be very slow this mirror would not have changed its position appreciably in the interval required for the light to travel from M to R and back. The light reflected from M would therefore pass through the lens system and produce an image at P. This image would remain as long as light was incident upon R. When this condition was no longer true there was no image at P and the field appeared dark. Thus, as long as the

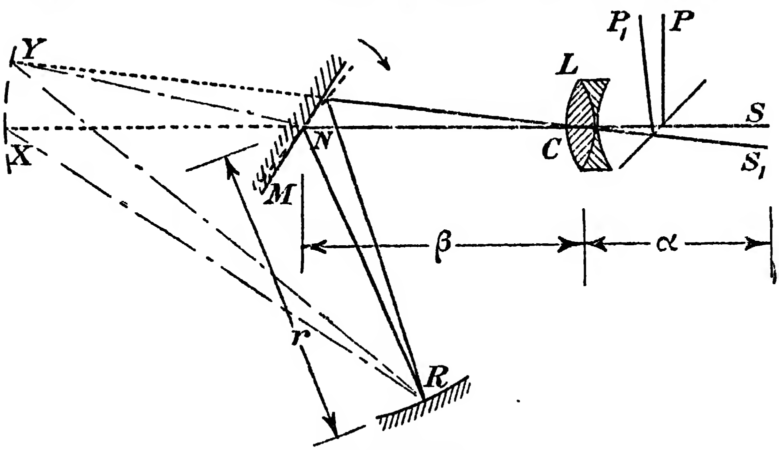


FIG. 27·3 (b).

speed of rotation was slow, brightness and darkness followed in turn at P. When the speed was increased so that the alternations occurred more than ten times a second a permanent image remained at P owing to the persistence of the visual impression on the retina. The brightness of this image would be reduced in the ratio of the arc R to the circumference of the circle of which this was part. When, however, the speed of the mirror M was increased so that it had moved through a small angle in the interval of time required by the light to travel from M to R and back again the light reflected from M did not retrace its original path but one which, if M rotated in a clockwise direction, caused the final image to be formed at P_1. From the observed shift PP_1 and the speed of rotation the velocity of light was calculated as follows. The diverging beam of light falling upon L would converge to a point image at X, Fig. 27·3 (b), in the absence of the mirror M. Actually it formed an

image at R a point on the concave mirror, the two points X and R being on a common normal to the mirror M and at equal distances from its plane, i.e. X was the image of R. When the mirror M was rotated about a vertical axis through N the image of R was at Y, let us say, where X, Y, and R lie on a circle whose centre is at N. By joining Y to the centre C of the lens L and producing it such that $CS = CS_1$ the position of the image was obtained when the glass plate was removed. When this plate was present the displaced image was at P_1 where $PP_1 = SS_1$.

Let $\alpha = CS$, $\beta = CN$, $r = NR$, $c =$ the velocity of light and Δ the displacement when the mirror M made n revolutions per second. Then the time to travel the distance $2NR = \frac{2r}{c}$ and in this time the rotating mirror has turned through an angle $2\pi n \cdot \frac{2r}{c} = \frac{4\pi nr}{c}$ radians. Now XR and YR are perpendicular to the two positions of the mirror M and therefore $X\widehat{R}Y =$ angle of rotation; hence, since $NR = NX = YN$, from two sets of congruent triangles X, Y, and R are on the circumference of a circle with N as centre, so that $X\widehat{N}Y = 2X\widehat{R}Y =$ twice the angle of rotation.

$$\therefore \qquad XY = 2r \cdot \frac{4\pi nr}{c} = \frac{8\pi nr^2}{c}.$$

$$\text{But } PP_1 = SS_1 = \frac{\alpha}{(\beta + r)} XY.$$

Therefore

$$\Delta = PP_1 = \frac{8\pi nr^2\alpha}{c(\beta + r)}, \text{ or } c = \frac{8\pi nr^2\alpha}{\Delta(\beta + r)} = \frac{8\pi nr\alpha}{\Delta}, \text{ if } \frac{\beta}{r} \to 0.$$

In Foucault's experiment $n = 800$ rev.sec.$^{-1}$, $r = 4$ metres, but the light was made to traverse this distance five times, so that in effect

$r = 20$ metres, $\Delta = 0{\cdot}7$ mm. which gave $c = 2{\cdot}98 \times 10^{10}$ cm. sec.$^{-1}$. Foucault also employed this method to determine the velocity of light in water. A tube containing water was placed between N and R, and a converging lens was placed in front of this tube to compensate for the effect of the water on the convergence of the rays passing through it. The displacement of the image was greater than before. From this Foucault concluded 'La lumière se meut plus vite dans l'air que dans l'eau,' but he made no estimate of the ratio of the two velocities. In 1883 Nicholson, using an apparatus similar to Foucault's, measured the two displacements and found the ratio to be 1·330. This agreed well with the refractive index of water, 1·334, as required by the wave theory which is briefly outlined in the next chapter.

Michelson's Experiments.—Foucault's method for determining the velocity of light suffers from the fact that the brightness of the image decreases as the distance RM increases. MICHELSON showed that, if the lens were placed at a suitable point between M and R, the brightness was independent of the distance RM provided that a lens of sufficiently long focal length was used. In his first experiments this distance was augmented to 600 metres and a shift of 133 mm. obtained. In 1926 Michelson published an account of a new and somewhat modified method. The two stations were Mount Wilson and Mount San Antonio in California, their distance apart being about 22 miles and the time for light to go and return about 0·00023 second. He found the velocity to be 299796 km. sec.$^{-1}$ *in vacuo* with an error of about 1 in 100,000.

EXAMPLE XXVII

Describe Foucault's method for measuring the velocity of light in air and in water, and discuss the importance of the results he obtained.

CHAPTER XXVIII

THE EMISSION AND WAVE THEORIES OF LIGHT

Newton's Corpuscular or Emission Theory.—According to Newton the sensation of light was due to the mechanical impact of swarms of small particles emitted by the luminous object observed. He assumed that they travelled in straight lines except when they approached infinitely close to matter, when their rectilinear paths became modified. The phenomena of reflexion and refraction were attributed to the modifications thus introduced. To explain why some particles were reflected while others were refracted, Newton assumed that they were subject to 'fits' of easy

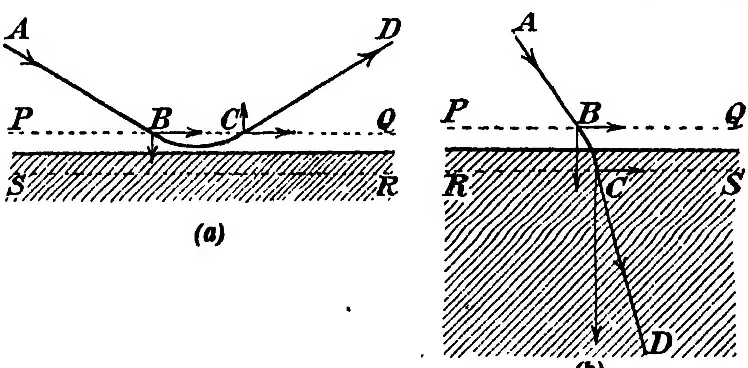


FIG. 28·1.—Reflexion and Refraction of a Light Corpuscle.

reflexion and of easy transmission. To explain the phenomenon of reflexion Newton assumed that, if during the time when one of these corpuscles was in a 'fit of easy reflexion' it approached very close to the interface of two media, the component of its velocity perpendicular to the interface began to experience a repulsive action. The component parallel to the surface was not altered. As soon as the normal component of the velocity began to change, the path of the particle became curved; eventually, when this component had become zero, the path was momentarily parallel to the tangent to the interface at the 'point of contact.' The repul-

sive action was still operative, however, and the particle therefore moved away from the surface along a curved path until its normal component of velocity experienced no further change: this happened when this component had reached its initial value numerically but with sign reversed. The particle would then emerge free from the influence of the interface and travel in a straight line with a velocity whose components were numerically as before, but with the normal one changed in direction. The path of the reflected particle must therefore be inclined to the normal to the surface, at the point where the normal component of the velocity vanished, at an angle equal to that made by the incident ray with the same normal. These facts are made more clear by referring to Fig. 28·1 (*a*). Since there had been no component of velocity normal to the plane of the diagram, the motion must be wholly in that plane. Thus the laws of reflexion are wholly accounted for by the corpuscular theory of light.

On the other hand, when a particle having a 'fit of easy transmission' came under the influence of the interface, i.e. nearer to the surface than PQ, Fig. 28·1 (*b*), the normal component of its velocity was increased and even after entering the medium this component continued to increase until the corpuscle was beyond the region of influence of the interface, i.e. beyond RS. It then travelled through the medium with its normal component of velocity increased but having the same component of velocity parallel to the surface. Its resultant velocity was therefore increased. If the first and second media are air and glass respectively, say, let c_1 and c_2 denote the velocities of light in the two media, while θ_1 and θ_2 are the angles of incidence and refraction. Since the components of the two velocities parallel to the surface are identical we have

$$c_1 \sin \theta_1 = c_2 \sin \theta_2$$

i.e.

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{c_2}{c_1} = \mu,$$

where μ is the refractive index of glass. For the two media considered $\mu > 1$, so that *if* this theory is true $c_2 > c_1$. Foucault and others have measured the velocity of light in various media by placing a rod of the material or a tube containing it, if it were a liquid, between M and R [cf. Fig. 27·3 (*a*), p. 537]. They found that the velocity of light in all material media was less than that in air. Hence the corpuscular theory of light cannot be valid.

Wave Theory and the Principle of Huyghens.—In 1678 a Dutchman named Huyghens postulated that light was a wave motion propagated in the æther, the æther being an all-pervading medium in which matter exists. This æther was originally

supposed to possess density and elasticity since it acquired kinetic energy when set in motion and potential energy when it was strained. Huyghens assumed that the properties of the æther were the same in all directions, i.e. it is isotropic. To account for the propagation of waves let us assume that O, Fig. 28·2, is the centre of a spherical disturbance which at the instant considered has reached AB. It must be remembered that the disturbance is really spread over the surface of a sphere with centre O so that AB is merely a portion of the circle which is the intersection of the sphere with the plane of the paper. AB is therefore a portion of the wave front due to a source at O. Huyghens imagined that each point on a wave front is the centre of a new disturbance known as a secondary wave. At a time t after the disturbance has reached AB each of these secondary disturbances will have wave-fronts extending over spheres of radii ct, where c is the velocity of propagation. The resultant wave front will be the envelope of these, i.e. it will be represented by A′B′. This is the arc of a circle with its centre at O.

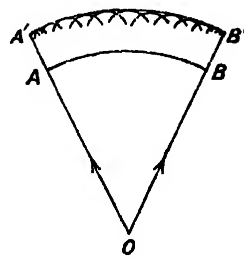


FIG. 28·2.

Newton appreciated all these points in the wave theory, but he could not reconcile it with the fact that all waves known to him (sound and water waves) could bend round corners, whereas light was propagated in straight lines. It is now known that the amount of bending depends directly upon the wave-length of the disturbance, and it is only because the wave-length of light is so small that its bending round corners is so minute that we are justified in regarding the propagation of light as being approximately rectilinear. FRESNEL showed that the wave theory would account for this approximately rectilinear propagation, but his arguments are beyond the scope of this book. However, let us see how the wave theory accounts for the reflexion and refraction of light.

Reflexion of Plane Waves.—Let CE, Fig. 28·3, be an interface between two media, and ABC the trace of a plane wave-front striking the interface at C. Then a, b, and c, normals to the wave-front, are 'rays of light.' According to the Principle of Huyghens, secondary wavelets are immediately formed when the wave-front reaches CE, so that every point in CE becomes in turn the centre of a secondary disturbance sent back into the first medium. The first point in CE to send out such waves is C while the last is E. When the disturbance reaches E the secondary wavelet from C will have moved forward a distance ct where c is the velocity in the first medium and t the interval of time between the arrival of the incident

wave at C and E respectively. From E draw EK a tangent to the wavelet from C. We have to show that this tangent is the trace of the reflected wave-front. The right-angled triangles ACE and KCE are congruent, for CE is common to both, and AE = CK since each is equal to ct; $\therefore \widehat{ACE} = \widehat{KEC}$. Now in the absence of the interface CE the wave-front ABC would have reached the position EGH in time t. Hence the $\triangle$'s ACE and CEH are congruent. $\therefore \widehat{ACE} = \widehat{CEH}$. Also the ray b which meets the surface at F would similarly have arrived at G, FG being perpendicular

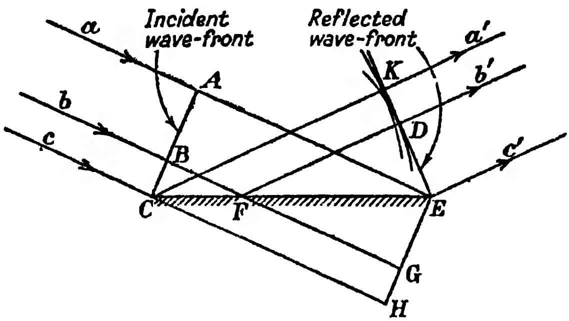


FIG. 28·3.—Reflexion of Plane Waves at a Plane Surface.

to EGH. Through F draw FD normal to KE to meet it in D. The $\triangle$'s FDE and FGE are congruent, for FE is common to both and $\widehat{DEF} = \widehat{GEF}$, since each is equal to $\widehat{ACE}$, as proved above. Hence FD = FG. But the radius of the secondary wave from F is FG; thus, the disturbance from F touches the line EK at D. In the same way it may be shown to be true for every wavelet originating from points in CE. Thus EK is the reflected wave-front. Also, by definition, $\widehat{ACE}$ is the angle of incidence, and $\widehat{CEK}$ the angle of reflexion, and from the above it follows that these are equal.

Reflexion of Spherical Waves at a Plane Surface.—Let O, Fig. 28·4, be a source of spherical waves situated in front of a plane mirror represented by CD. Consider a spherical wave which in the absence of the mirror would have attained a position AEB. This wave actually touches the mirror first at P where OP is normal to CD. By the time it has reached C and D, so that these are just becoming the centres of two secondary disturbances sent back into the first medium, the secondary wave from P will have a radius PF and the secondary disturbance from another point R will have

a radius RS where RS = RT, T being the intersection of OR produced and the arc AEB. It is clear that the envelope of all these secondary disturbances, which is the reflected wave-front, will be the arc of a circle identical with the arc CED, for this arc may be considered as the envelope of circles having their centres on CD and radii equal to the distance of their centres from the arc measured along the normals through their centres to the arc. The centre of the reflected wave-front is I where IF = EO, i.e. IP = OP.

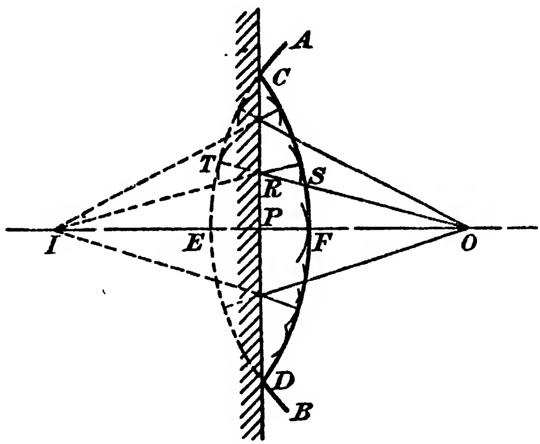


FIG. 28·4.—Reflexion of a Spherical Wave at a Plane Surface.

LEMMA. ***The curvature of a small circular arc:*** Let AOB, Fig. 29·10 (*b*) [p. 564], be a small arc of a circle whose radius is *r*. Let M be the mid-point of the chord AB. Then OM is termed the ***sagitta*** of the arc AOB. It is well known that $AM^2 = OM \, . \, MD$. When OM is small, MD becomes $2r$, so that $OM = \frac{AM^2}{2r}$. Since the curvature of a circular arc is defined as the ratio $\frac{\textit{angle}}{\textit{arc}}$ and this is $\frac{1}{r}$, it follows that for spherical arcs on a common chord AB, for example, the sagitta OM is directly proportional to the curvature of the corresponding arc AOB.

Reflexion of a Spherical Wave at a Concave Spherical Surface.—Let O, Fig. 28·5, be a point source of waves situated on the principal axis of a concave mirror whose centre of curvature is at C. Let APB be a section of this mirror in the plane of the paper. A spherical wave, whose trace is AQ_1B, diverging from O first meets the mirror at A and B. Let this be the instant from

which time is measured. Then by the time that the point Q_1 has reached the mirror at P, the secondary wavelets which originated at A and B at zero time will have extended so that their radii are equal to PQ_1. If $A_2Q_2B_2$ is an intermediate position of the wave diverging from O and under discussion, then when Q_2 reaches P, the secondary wavelets from A_2 and B_2 will have radii PQ_2. The reflected wave will be the envelope of all such secondary wavelets. Now let us assume that we have only a mirror of small aperture. Then we may assume that the reflected wave-front is spherical, its curvature being different from that of the incident wave. Moreover, if AK is drawn parallel to PC to meet the reflected wave-front in K, then AK will be, in the limiting case considered, equal to

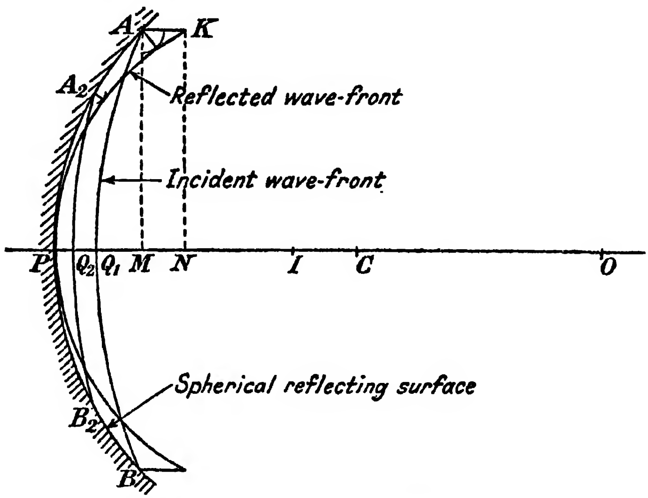


FIG. 28·5.—Refraction of Spherical Waves at a Spherical Surface.

the radius of the wavelet from A, viz. to PQ_1. Through A and K draw AM and KN normal to CP. Then the curvatures of the incident wave, the mirror and the reflected wave are proportional to the sagittæ Q_1M, PM, and PN respectively since the chords of the corresponding arcs are equal. Now

$$PN = PM + MN = PM + PQ_1 = 2PM - Q_1M,$$

or $$PN + Q_1M = 2PM,$$

i.e. the curvature of the reflected wave plus that of the incident wave is twice that of the mirror. If I is the centre of curvature of the reflected wave, with the usual notation, we have

$$\frac{1}{v} + \frac{1}{u} = \frac{2}{r}.$$

Refraction Plane Waves at a Plane Surface.—Let MN, Fig. 28·6, be a plane interface between two homogeneous and isotropic media in which light travels with velocities c_1 and c_2 respectively. Suppose that AB is the trace in the plane of the diagram of a plane wave whose direction of propagation is also in that plane. The disturbance in the second (or lower) medium first originates at A. Let this be the instant from which time is measured. Let t be the time required for the portion of the wave-front at B to reach E, where E is on the interface and BE is normal to AB. Then at time t a secondary wavelet will just be originating at E, while the secondary wavelet from A will have acquired a radius c_2t, or AC (say). Through E draw that tangent plane to

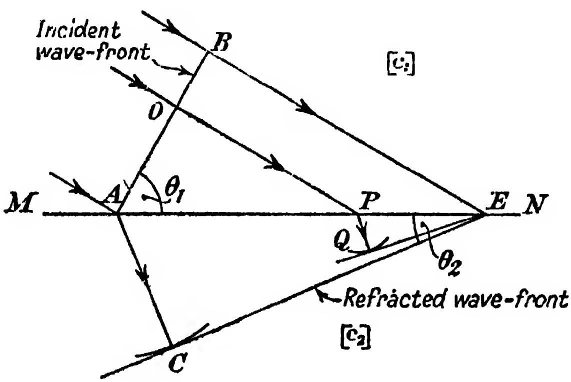


FIG. 28·6.—Refraction of Plane Waves at a Plane Surface.

the sphere whose centre is A and radius AC. If EC is the trace of this plane we have to show that EC is the refracted wave-front. To do this suppose that the point O in the wave-front AB reaches P, a point in AE, at time t', OP being normal to AB. Then when B reaches E, the secondary wavelet from P at the time t, will have a radius $c_2(t - t')$. If EQ is the tangent from E to the wavelet, $PQ = c_2(t - t')$. Hence

$$\frac{PQ}{AC} = 1 - \frac{t'}{t} = 1 - \frac{AP}{AE} = \frac{PE}{AE},$$

i.e.

$$\frac{PQ}{PE} = \frac{AC}{AE},$$

or, $\sin \widehat{AEC} = \sin \widehat{PEQ}$, so that EQ and EC must coincide. In this way it may be shown that all the secondary wavelets from points in the interface will touch CE, so that this is the refracted

wave-front. Moreover, if θ_1 and θ_2 are the angles indicated, they are respectively the angles of incidence and of refraction, so that

$$\frac{\sin\theta_1}{\sin\theta_2}=\frac{\mathrm{BE}}{\mathrm{AC}}=\frac{c_1}{c_2}.$$

Thus the constancy of the ratio of the sine of the angle of incidence to the sine of the angle of refraction is established by the wave-theory of light. This ratio is termed the refractive index of the second medium with respect to the first, and is denoted by ${}_1\mu_2$.

Let c_0 be the velocity of light in a vacuum, and let μ_1 and μ_2 be the refractive indices of the two media referred to a vacuum. Then

$$\frac{\sin\theta_1}{\sin\theta_2}=\frac{c_0/c_2}{c_0/c_1}=\frac{\mu_2}{\mu_1},$$

which gives

$$\mu_1\sin\theta_1=\mu_2\sin\theta_2.$$

Let us now suppose that $\theta_2 < \theta_1$, so that the lower medium is the more highly refracting, i.e. the refracted ray is bent towards the normal at the point of incidence. This occurs, for example, when light passes from water to glass. The experiments of Foucault and of Michelson showed that the velocity of light in water was greater than that in glass, and that

$$\frac{c_{\text{water}}}{c_{\text{glass}}}={}_1\mu_2.$$

Thus the wave-theory, in the middle of the last century, received almost universal acceptance for, as already cited, the corpuscular theory could only account for the laws of refraction on the assumption that the velocity of light in a vacuum was less than its velocity in any other medium, a result which was now shown to be contrary to experimental evidence, whereas the wave-theory showed that the velocity of light in a vacuum must be greater than that in any other medium, a fact which experiment now confirmed. This long-standing conflict ended, the science of optics made rapid advances: at first it was assumed that light vibrations took place in an elastic solid æther, but later on it was shown by MAXWELL that the vibrations were electromagnetic.

Refraction of a Spherical Wave at a Spherical Surface.—Let C, Fig. 28·7, be the centre of curvature of a spherical surface between two homogeneous and isotropic media in which light travels with velocities c_1 and c_2 respectively. Let O be a point source in the first medium and on the principal axis of the interface. Then a spherical wave diverging from O is in such a position that its trace in the plane of the paper is $\mathrm{A_1PB_1}$ when it first meets the

interface at P. Spherical wavelets then travel from P into the second medium. Suppose that when A_1 reaches A, the first wavelet from P has acquired a radius PQ_2. The refracted wave-front will be the envelope of all the wavelets from every point on the interface, each wavelet beginning as the appropriate point on the incident wave reaches the interface. If we limit our discussion to that portion of the interface in the immediate neighbourhood of P, we may assume that the refracted wave-front is that portion of a sphere passing through A, Q_2, and B, its centre being at I on the principal axis of the system. In the absence of a second medium the wave from O and under discussion would have reached a position indicated by AQ_1B. Since the time of propagation from A_1 to A, or P to Q_1, in the absence of the second medium, is equal to that from P to Q_2 when this medium is present, we have

$$\frac{PQ_2}{c_2} = \frac{PQ_1}{c_1}.$$

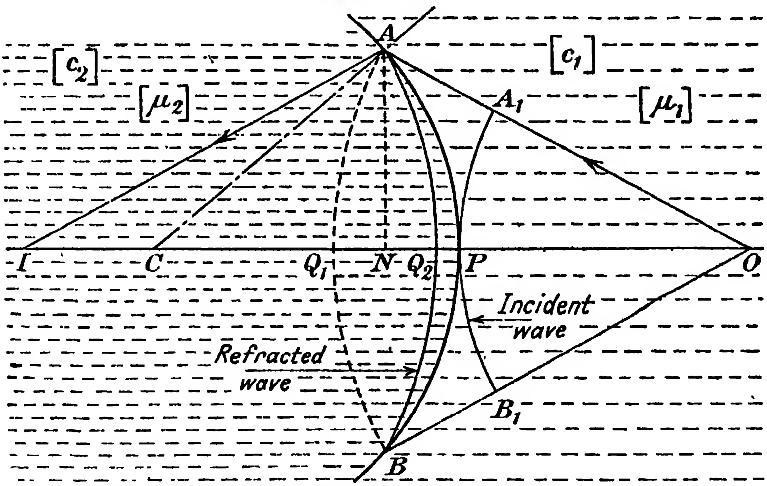


FIG. 28·7.—Refraction of Spherical Waves at a Spherical Interface.

Now the curvatures of the incident wave, the interface, and the refracted wave are respectively proportional to Q_1N, PN, and Q_2N, where N is the projection of A on OC. We have

$$NQ_2 + Q_2P = NP,$$
$$Q_1N + NP = Q_1P.$$

$$\therefore \frac{NQ_2 - NP}{c_2} = -\frac{Q_2P}{c_2} = \frac{PQ_2}{c_2}$$
$$= \frac{PQ_1}{c_1} = -\frac{Q_1P}{c_1} = \frac{-Q_1N - NP}{c_1}$$

$$\therefore \frac{-Q_2N + PN}{c_2} = \frac{-Q_1N - NP}{c_1} = \frac{-Q_1N + PN}{c_1}$$

$$\therefore \frac{-\frac{1}{v} + \frac{1}{r}}{c_2} = \frac{-\frac{1}{u} + \frac{1}{r}}{c_1}$$

Multiplying throughout by c_0, the velocity of light *in vacuo*, and remembering that μ_1 and μ_2, the absolute refractive indices of the two media, are respectively $\frac{c_0}{c_1}$ and $\frac{c_0}{c_2}$, we have

$$\mu_2\left[\frac{1}{r} - \frac{1}{v}\right] = \mu_1\left[\frac{1}{r} - \frac{1}{u}\right],$$

which is identical with the appropriate equation obtained from geometrical optics.

Refraction by a Lens.—Let O, Fig. 28·8, be a point source of light situated on the principal axis of a converging lens. Then a wave-front, occupying the position CLD in the absence of the lens, will have its central portions retarded so that CMD (full line) is the

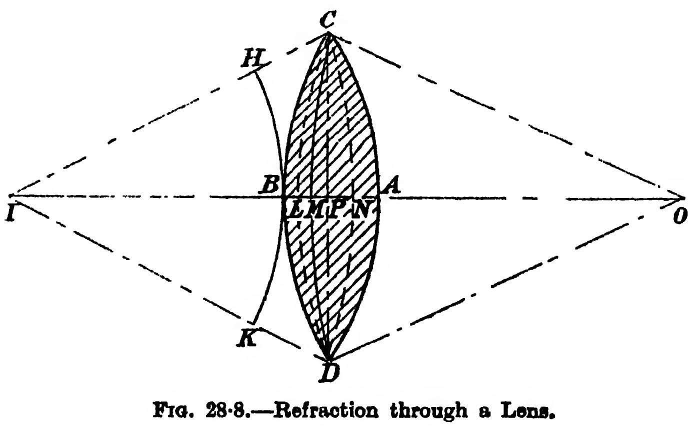


Fig. 28·8.—Refraction through a Lens.

actual wave-front in the lens. The wave-front emerging from the lens will have its central portions still further retarded with respect to its outer ones. Let us assume that the wave-front on first emerging has the form HBK, which is the arc of a circle when the distance AB is small. Let I be the point to which the emergent waves converge. [There are cases in which I is on the same side of the lens as O—the emergent waves then diverge from this point and the image is virtual.] With centre I and radius IC describe an arc to cut the axis in N. Then the time for the disturbance to travel

in air from O to I via C must be equal to the time required for it to travel from O to I via A and B, i.e. partly in air and partly in glass. Let V_a and V_g be the velocities of light in the two media. Then

$$\frac{OC}{V_a}+\frac{CI}{V_a}=\frac{OA}{V_a}+\frac{AB}{V_g}+\frac{BI}{V_a},$$

i.e.
$$OC+CI=OA+\mu AB+BI,$$
$$OL+NI=OA+\mu AB+BI$$
$$AL+NB=\mu AB$$
$$=\mu(AP+PB),$$

where CPD is normal to OI,

i.e. $(LP+PA)+(NP+BP)=\mu\,(AP+PB)$.

Consequently,

$$\frac{1}{|u|}+\frac{1}{|r_1|}+\frac{1}{|v|}+\frac{1}{|r_2|}=\mu\left[\frac{1}{|r_1|}+\frac{1}{|r_2|}\right],$$

$$\text{i.e. } \frac{1}{|v|}+\frac{1}{|u|}=(\mu-1)\left[\frac{1}{|r_1|}+\frac{1}{|r_2|}\right].$$

Introducing the usual convention for signs, we have,

$$\frac{1}{v}-\frac{1}{u}=(\mu-1)\left[\frac{1}{r_1}-\frac{1}{r_2}\right].$$

EXAMPLES XXVIII

1.—What are the principal characteristics of wave motion? By considering the refraction of a plane wave at a plane surface, show how the laws of refraction of light can be explained on the wave theory.

2.—Describe an experiment to show that light consists of waves. What are the essential differences between sound and light waves?

CHAPTER XXIX

INTERFERENCE AND DIFFRACTION

The Principle of Superposition.—The reason why Newton and his school were not convinced by the exponents of the wave theory was that if light were a vibratory motion in the æther, then how could its rectilinear propagation and the formation of shadows be explained? It was not until YOUNG, in 1801, introduced the principle of interference into optics that an explanation of the above in terms of the wave theory was forthcoming.

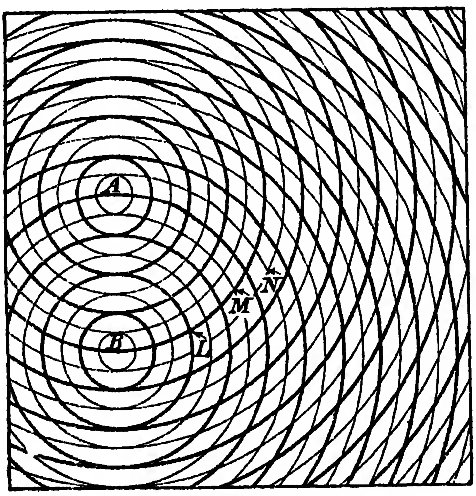


FIG. 29.1.

If two wave trains having the same wave-length travel across a mercury or water surface, each will produce the same effect as if it alone were present. The resultant displacement, at any instant, of a particle on the surface is therefore the algebraic sum of the displacements it would have at that same instant due to each separate train. Let us assume that A, Fig. 29·1, is the point where a style attached to a tuning-fork just touches the surface of some mercury. When the fork vibrates waves will originate at A and travel across

the surface. At any instant these waves may have crests represented by the thick line circles, and troughs by the others. If a second style is attached to the same prong and touches the mercury at B, a crest will originate at B simultaneously with one at A. The crests and troughs forming this second train are also indicated. At points, such as L, where a crest of one train coincides with a crest from the other the resulting displacement will be a maximum, while at points such as M where a crest of one train meets a trough in the other the resultant displacement will be zero. Also at N, and other similar positions, where trough meets trough, the displacement will be a maximum, only in a direction opposite to that at L. The two wave trains are said to have ***interfered*** at the points where the disturbance is zero and a stationary pattern will have been produced on the mercury surface.

If light consists of waves it should be possible to obtain interference patterns if two sources emitting waves of the same wave-length and in the same phase can be procured. Now every attempt to obtain interference with two different light sources, even though they are monochromatic, must necessarily fail, since the light vibrations from any source undergo rapid and abrupt changes in phase. To obtain permanent interference patterns the two sources must be either a real source and its image or else two images of the same source, for then any change of phase in the real source will cause a simultaneous and equal change in its image. The rays from such sources are said to be ***coherent.***

Investigation of the Combined Effects due to Two Separate Coherent Light Sources.—Let S_1 and S_2, Fig. 29·2 (*a*) be two such point sources of light, i.e. the amplitudes, phases, and wave-lengths, λ, are identical. Suppose that MN is a screen parallel

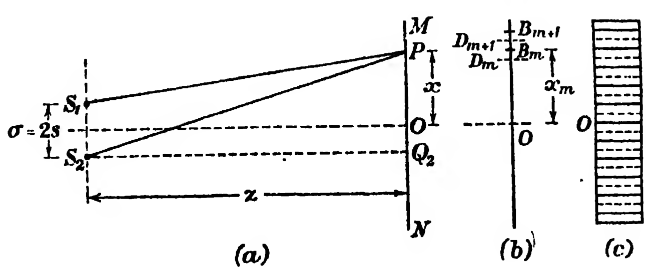


Fig. 29·2.

to S_1S_2 and at a perpendicular distance z away. Let the perpendicular bisector of S_1S_2 meet MN in O: then at O there will be brightness since the distances S_1O and S_2O are equal. To ascertain whether or not a bright fringe exists at P, a point in MN at a distance x

from O, it is necessary to discover whether or not $S_2P - S_1P$ is equal to an even or odd multiple of $\frac{\lambda}{2}$. Draw S_2Q_2 normal to MN. Then

$$S_2P - S_1P = [z^2 + (x+s)^2]^{\frac{1}{2}} - [z^2 + (x-s)^2]^{\frac{1}{2}} \qquad [\text{if } \sigma = 2s = 2.OQ_2]$$

$$= z\left[1 + \left(\frac{x+s}{z}\right)^2\right]^{\frac{1}{2}} - z\left[1 + \left(\frac{x-s}{z}\right)^2\right]^{\frac{1}{2}}$$

$$= z\left[1 + \frac{1}{2}\left(\frac{x+s}{z}\right)^2 - 1 - \frac{1}{2}\left(\frac{x-s}{z}\right)^2\right],$$

since $\left(\frac{x+s}{z}\right)^2$ and $\left(\frac{x-s}{z}\right)^2$ are small, so that we may use the well-known fact that if α is small $(1+\alpha)^{\frac{1}{2}} = 1 + \frac{1}{2}\alpha$.

$$\therefore \quad S_2P - S_1P = z\left[\frac{xs}{z^2} + \frac{xs}{z^2}\right]$$

$$= \frac{2xs}{z} = \frac{\sigma x}{z}.$$

Hence there will be a bright fringe at P if

$$\frac{\sigma x}{z} = 2m.\frac{\lambda}{2}, \qquad \text{where } m = 0, 1, 2, 3, \ldots$$
$$= m\lambda.$$

[m is called the *order of the fringe*.]

Similarly, there will be darkness if

$$\frac{\sigma x}{z} = (2m+1)\frac{\lambda}{2}.$$

Let B_m, Fig. 29·2 (*b*), be the position of the mth bright fringe, B_{m+1} that of the $(m+1)$th. Call $OB_m = x_m$; $OB_{m+1} = x_{m+1}$. Then

$$\frac{\sigma x_m}{z} = m\lambda; \quad \frac{\sigma x_{m+1}}{z} = (m+1)\lambda.$$

$$\therefore x_{m+1} - x_m = \frac{\lambda z}{\sigma},$$

i.e. the distance between consecutive bright fringes, for a given disposition of screen and light sources, is constant and equal to $\frac{\lambda z}{\sigma}$. Similarly, it may be shown that the dark fringes, D_m, D_{m+1}, etc., are equally spaced between the bright fringes.

Thus, if interference fringes can be produced and the fringe-width measured, it becomes possible to determine the wave-length

of the light used. Fig. 29·02 (*c*) is an attempt to indicate the nature of the interference pattern on the screen.

Fresnel's Biprism.—An optical system used by Fresnel, about 1820, to measure the wave-length of monochromatic light from observations on an interference pattern, makes use of a so-called biprism. ABC, Fig. 29·3, is a section of such a prism, the section being cut normal to the base of the prism. The prism is made of glass and the angle at A is very obtuse. Let S be a monochromatic source of light situated on the perpendicular bisector of the base BC and at a distance *a* from the prism. An eye placed near to the axis SA and on the side of the prism remote from S will see two virtual images S_1 and S_2. S_1 is produced by light waves from S incident upon AB and refracted through the prism; S_2 is formed similarly. The virtual images act as coherent sources. Thus, between the prism and the screen MN there are two refracted beams which overlap in the region HAK so that here interference fringes are obtainable. If *b* is distance between the prism and screen, $(a + b) = z$, and the fringe-width is therefore

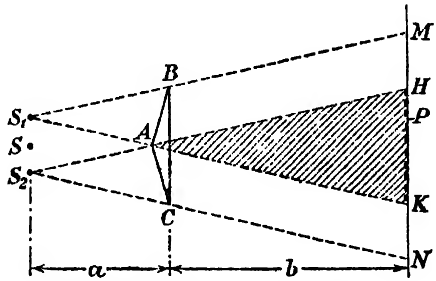


FIG. 29·3.—Fresnel's Biprism.

$$\frac{z\lambda}{\sigma} = \frac{(a + b)\lambda}{\sigma}.$$

This system of fringes may be observed in a low power microscope focused on the plane HK. In addition to the true interference fringes which are equidistant, another set of fringes will be seen superposed on them. These are easily distinguished, however, for their distance apart is not constant. They are diffraction fringes.

In actual practice the source of light is a very narrow slit parallel to the refracting edge of the prism. If the slit is widened or rotated about a horizontal axis, the biprism remaining fixed, the fringes disappear.

Investigation of the Effect of placing a Thin Plate in the Path of One of Two Coherent Light Beams.—Suppose that when a thin glass plate DE, Fig. 29·4, of thickness *t*, is placed so that the light from S_1, one of a pair of coherent sources, passes through it, the central fringe is displaced a distance *x* from O to P. Then the time for a disturbance to travel from S_1 to P must equal

that for a disturbance to travel from S_2 to P. If μ is the refractive index of the material of the thin plate, c_0 the velocity of light in air, c its velocity in glass, the above equality in the times of transit becomes

$$\frac{S_1P - t}{c_0} + \frac{t}{c} = \frac{S_2P}{c_0}.$$

Since $\mu c = c_0$, $S_2P - S_1P = (\mu - 1)t$.

But already it has been shown [cf. p. 554] that

$$S_2P - S_1P = \frac{\sigma x}{z}.$$

$$\therefore (\mu - 1)t = \frac{\sigma x}{z}.$$

Hence from observations on the shift of the central fringe, either $(\mu - 1)$ or t may be determined according as t or μ is known as

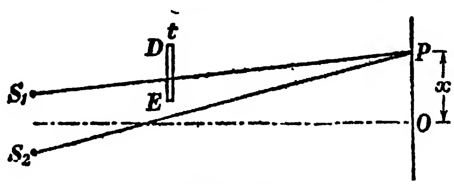


FIG. 29·4.

well as σ, x, and z. It is only possible to identify the central fringe if white light is used: for in such circumstances each coloured constituent produced its own fringe system with its own characteristic width, so that overlapping occurs to such an extent that only a few coloured fringes round the central white one are to be seen.

The Diffraction Grating.—A diffraction grating consists of a large number of equidistant and parallel lines ruled on a plate of glass or of speculum metal. For some purposes it is advantageous to have the surface concave and part of a cylinder, but the only type we shall consider is that in which the surface is plane. The dark short lines shown in Fig. 29·5, represent the opaque portions of a grating. Let us consider a system of parallel rays, OA being one of them, incident upon the grating. Let us further assume that after refraction at the front surface of the glass the ray OA meets B, a point at the extremity of one of the rulings on the grating. A secondary wavelet originates at B. Similarly wavelets proceed from the points E, G, H, etc.

From A and B draw AC and BF normal to the incident and diffracted rays PD and EN respectively. At A and C the phases of the incident waves are the same. At B and F any difference in phase

will be due to the different times for the disturbances to be propagated via the paths AB and CDEF. This difference of time is

$$\frac{\mathrm{CD}}{c_0}+\frac{\mathrm{DE}}{c}+\frac{\mathrm{EF}}{c_0}-\frac{\mathrm{AB}}{c}=\frac{\mathrm{CD}+\mathrm{EF}}{c_0}$$

since DE = AB, where c_0 and c are the velocities of light in air and glass respectively. The difference in path is therefore CD + EF. Now any difference in phase existing at B and F will

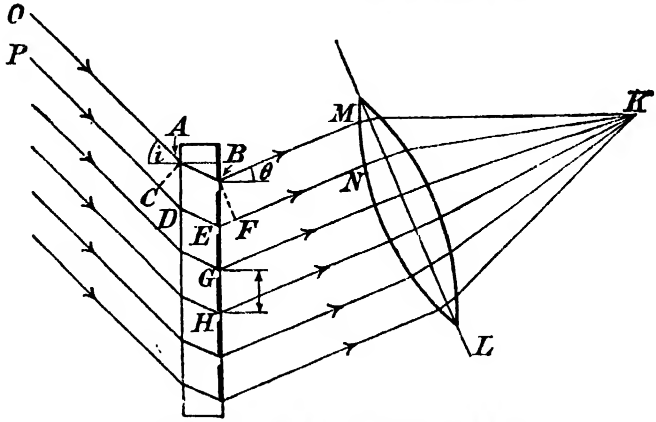


FIG. 29·5.—A Plane Diffracting Grating.

still exist at K since the time for the disturbances to travel from B and F to K is independent of the actual path taken. We therefore see that the two waves will reinforce each other at K if the path difference CD + EF is an even number of half wave-lengths, i.e. if

$$\mathrm{CD}+\mathrm{EF}=\frac{2m\,.\,\lambda}{2}=m\lambda$$

where m is an integer and λ the wave-length of the incident radiation. But CD + EF = AD sin i + BE sin θ = $(a + b)$ (sin i + sin θ), where $(a + b)$ = BE, the sum of the widths of the grating aperture and a space: it is equal to $\frac{1}{n}$ where n is the number of lines per unit length on the grating. Hence for reinforcement

$$(a+b)(\sin i+\sin\theta)=m\lambda.$$

In general it is usual to arrange the grating normal to the incident rays when sin i becomes zero, and we have

$$(a+b)\sin\theta=m\lambda.$$

For any given value of m, every wave-length λ present has a definite angle θ. Thus a spectrum of the light received is produced. If $m = 1$, we have a '***first order spectrum***': for $m = 2$, we have a '***second order spectrum,***' etc.; m may also be a negative integer.

The Visual Examination of a Spectrum Produced by means of a Diffracting Grating.—A typical arrangement is shown in Fig. 29·6. The slit S, illuminated by the light whose spectrum is required, lies at the first focus of a converging lens, L_1, and the grating, G, is arranged so that the parallel beam of light emerging from L_1 falls normally on the grating. Corresponding to each wave-length in the incident light there is a system of diffracted beams—one in each order. Here we shall only consider the first order. Such a beam is shown in the diagram, and it is brought to a focus at F, the second focus of the converging lens L_2. The diffracted image at F is examined visually by a lens L_3 arranged so that a beam of parallel light enters the eye.

[An identical spectrum appears on the other side of the axis of the collimator.]

The remarks made above with reference to the diffracted rays

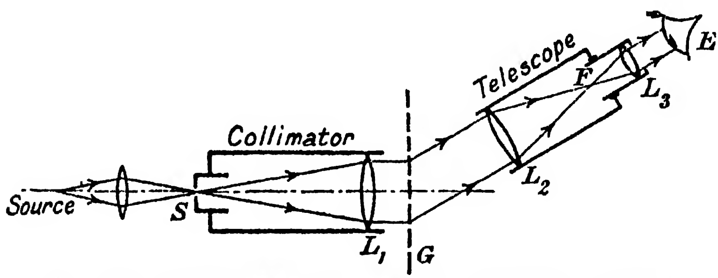


FIG. 29·6.—Spectrometer: its use with a Diffraction Grating.

from the pair of points B and E apply equally well to any such pair so that the above equations apply to the grating as a whole.

To measure the wave-length of some monochromatic light the grating is mounted on a spectrometer table normal to the light from the collimator, and the telescope having been focused for parallel light is turned to view the first diffracted image, corresponding to $m = 1$. Actually there will be two such images at equal angles on opposite sides of the normal to the grating. The angle between them is measured—it is 2θ. Knowing the number of lines per unit length of the grating the above formula enables λ to be calculated.

The second order spectrum is then looked for, but it may be too faint for purposes of accurate observation. If it can be located λ may again be calculated using $m = 2$.

If the grating is illuminated with white light a continuous spectrum will be obtained, but it may not be so pure as that obtained with prisms on account of the overlapping of spectra of different orders.

The Localization (Situation) of Interference Fringes.—Let S, Fig. 29·7 (*a*), be a ***point source*** of monochromatic light of wave-length λ, and let I be an apparatus—a so-called interferometer—for producing a path difference between coherent rays of light. Let P be a point in the space where interference may occur. Consider the two rays SACP and SBDP. Let Δ be the difference in optical path between these two rays. If Δ is an odd multiple of $\frac{\lambda}{2}$, darkness will result at P; brightness occurs when Δ is an even multiple of $\frac{\lambda}{2}$. Now when S and P are fixed, Δ is completely determined. If, therefore, a screen is held in the neighbourhood of P, the illumination on the screen will vary from point to point, darkness and brightness alternating. Such fringes are obtained wherever the screen is held, providing, of course, it is in the path of the interfering beams; such fringes are said to be ***non-localized.***

Extended Source: Localized Fringes.—Suppose now that S Fig. 29·5 (*b*), is an ***extended source*** of monochromatic light, i.e. it is a region containing many point sources. It will be found that interference fringes are to be seen only in definitely located positions—such fringes are said to be ***localized.*** Let SACP and SBDP be two coherent rays meeting at P and suppose that the eye is focused on P. The pupil subtends an appreciable angle at P: it therefore receives from P not only the above rays (in reality they are so near together that if drawn to scale they would coincide) but a conical pencil made

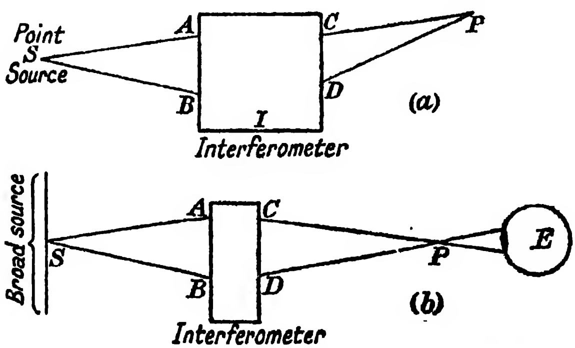


FIG. 29·7.—Localization of Interference Fringes.

up of similar pairs of rays coming to P from adjacent parts of the source via adjacent parts of the interferometer. In general, the value of δ for each such pair will be different and the illumination at P, which is due to all the light reaching the eye from that point, will neither be a maximum nor a minimum. The same applies to other points near P, so that the illumination is uniform and no fringes are visible. It may be shown, however, that the rate at which δ varies with the inclination of the rays passing through a given point is not the same at all distances from the interferometer, and at certain distances it has a minimum value. In this region fringes will be observed if the cone of rays utilized from each point of the field is

not too wide. [Fringes may be seen which cannot be photographed on account of the wider angle of the cone of light received by a camera lens.]

Interference Fringes obtained with Thin Parallel Plates in Monochromatic Light.—Consider a ray of monochromatic light, OA, Fig. 29·8, incident at an angle α upon the upper surface of a thin parallel plate of thickness, t (a film of air enclosed between two parallel glass surfaces is a 'plate' in so far as this discussion is concerned, although we shall, for the sake of being definite, consider a plate of, say, glass). This gives rise to a reflected ray, AH and a refracted ray, AB, which is reflected at the rear surface of the film—there is also a refracted ray proceeding from this point, but it does not concern us here. At C there is produced a reflected ray, CD, and a ray, CK, emerging from the film in a direction parallel to AH. Similarly, other reflected and refracted rays are produced.

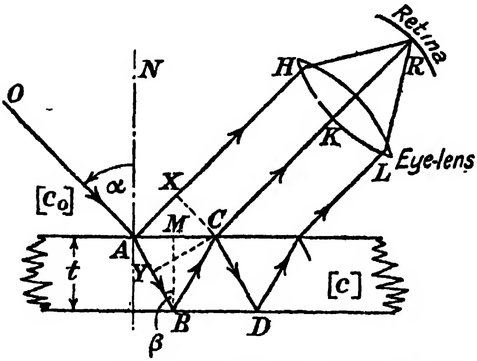


Fig. 29·8.—Interference Fringes with a thin parallel plate (reflected light).

By considering any two adjacent rays of the system of parallel rays emerging from the plate it may be shown that there is a constant path difference between them. Let us select the rays AH, CK. Here it must be emphasized that the intensity of the pattern formed on the retina of the eye receiving the rays would involve a discussion of all the light rays reflected from the surface of the film and entering the eye, but their effect will not greatly modify the result obtained by considering only two rays, since the intensity of the rays neglected is much less than that of the rays AH and CK.

Let CX and CY be drawn normal to AH and AB respectively. Consider the difference in time required for the light to reach the retina when it travels along OAHR and OABCKR. The time of travel along OA may be neglected since it is common to each ray; also the times from X to R and C to R, since they are equal—a fundamental property of any lens system. Let c_0 and c be the velocities of light in air and in the medium of the plate. Let λ be the wave-length of the light, this being measured in air.

The time for the light to travel from A to X is $\dfrac{AX}{c_0}$.

" " " " A to C via B is $\dfrac{AB + BC}{c}$

∴ the second ray is apparently delayed by an amount

$$\frac{AB + BC}{c} - \frac{AX}{c_0}$$

But $\frac{AX}{c_0} = \frac{AY}{c}$, since the time of travel from A to X is equal to that from A to Y [cf. p. 547].

$\therefore$ Apparent time difference is $\frac{AY + YB + BC}{c} - \frac{AX}{c_0}$

$$= \frac{YB + BC}{c} = \frac{2t \cos \beta}{c},$$

where β is the angle of incidence on the lower face.

The above time difference is equivalent to a path difference measured in air of an amount

$$\frac{2t \cos \beta}{c} \cdot c_0 = 2\mu t \cos \beta.$$

The word 'apparent' has been inserted in the above argument since an important correction has to be made. The electromagnetic theory of light establishes the fact that when light is reflected at a medium-air interface there is no change in phase: when the reflexion occurs at an air-medium interface a phase difference of π, equivalent to a path difference of $\frac{1}{2}\lambda$, λ being the wave-length in air, is introduced. This fact has been established experimentally [cf. the corresponding instance in acoustics]. In the present instance this change in phase occurs at A. Hence the effective time of transit from the arrival of the disturbance at A until it reaches X is

$$\frac{AX + \frac{1}{2}\lambda}{c_0}.$$

The effective path difference is therefore

$$\left[\frac{AY + YB + BC}{c} - \frac{AX + \frac{1}{2}\lambda}{c_0}\right]c_0$$

$$= 2\mu t \cos \beta - \frac{\lambda}{2} = \delta \text{ (say)}.$$

For the rays to reinforce one another at R, the above expression must be equal to an even multiple of $\frac{\lambda}{2}$, even to $m\lambda$, where m is a positive integer, including zero, i.e.

$$2\mu t \cos \beta - \frac{\lambda}{2} = m\lambda$$

or $$2\mu t \cos \beta = (2m + 1)\frac{\lambda}{2}.$$

For darkness, the necessary condition is

$$2\mu t \cos \beta - \frac{\lambda}{2} = (2m + 1)\frac{\lambda}{2}$$

or $$2\mu t \cos \beta = m\lambda.$$

The light transmitted through the plate will also exhibit interference effects—in fact, the system of fringes will be such that where the intensity is a minimum in one, it is a maximum in the other. The intensity of the bright fringes will be the same in each instance, if no absorption occurs in the plate, but the minima in the transmitted light are never well defined since the light reflected back into the

plate is only a small fraction of that which is transmitted. A calculation of the retardation with transmitted light is instructive and will therefore be given.

Fig. 29·9 indicates the formation of the rays of light to be considered. As before, we limit our discussion to two rays, viz. BK and

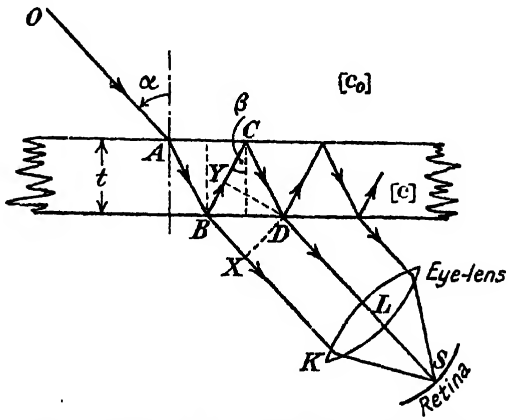


FIG. 29·9.—Interference Fringes with a thin parallel plate (refracted light).

BCDL. The difference in time of travel from O to S by the paths indicated is

$$\frac{AB + BC + CD}{c} - \left(\frac{AB}{c} + \frac{BX}{c_0}\right) = \frac{BY + YC + CD}{c} - \frac{BX}{c_0} = \frac{YC + CD}{c}.$$

The path difference measured in air is therefore

$$\frac{YC + CD}{c} \cdot c_0 = 2\mu t \cos \beta.$$

This path difference is the true path difference since there is no change of phase when a reflexion occurs at a medium-air interface. Hence, for darkness,

$$2\mu t \cos \beta = (m + \tfrac{1}{2})\lambda,$$

and for brightness

$$2\mu t \cos \beta = m\lambda.$$

The Necessity for a Broad Source.—Suppose a very thin soap film is made and sodium light, for example, from a point source reflected from it. Fringes only appear across the image of the source. If a broad source of such light is used, the whole film appears to be crossed by interference fringes. Since these fringes do not move relatively to the film when the eye of the observer moves, they must be located in the film. It may be shown that the fringes are formed at infinity when parallel light is used and that the fringes are then circular. The subject is, however, too difficult for us to discuss further.

The Colours of Thin Films.—Sufficiently thin films of transparent substances exhibit brilliant colours when examined in white light. Such colours are to be seen in soap bubbles, in thin oil films

resting on a surface of water, and in the thin oxide layers coating various metals. It has been shown above how interference effects may be obtained with thin films when monochromatic light is used to view them. With white light, when the condition for a minimum intensity is satisfied for one colour, it is not satisfied for others and if one colour is removed from white light the complementary colour appears.

Newton's Rings.—A thin film of air of slowly varying thickness is produced when a converging lens whose surfaces have large radii of curvature is placed on a flat surface [cf. Fig. 29·10 (*a*)]. If the point of contact is viewed in white light it will be seen surrounded by coloured rings. These were first observed by HOOKE in 1665 and their radii measured by NEWTON: YOUNG gave the first satisfactory explanation of them. The coloured rings are due to interference effects between light waves reflected at the upper and lower surfaces of the air film between the lens and plate. If monochromatic light is used many more rings are observed. Let us calculate the condition for the intensity to be a maximum or a minimum when the rings are viewed by reflected light.

We shall assume that there is good optical contact between the lower surface of the lens and the plate, and that the light falls normally on the plate—this latter condition is true only if r is large and the fringes limited to the region close to the point of contact of the surfaces. If R is the radius of curvature of the spherical surface, ρ that of a ring, and PN the thickness of the air film, then

$$\mathrm{PN} = \tfrac{1}{2}\frac{\rho^2}{R},$$

for if we imagine the circle to be completed as in Fig. 29·10 (*b*), then

$$\mathrm{PN} = \mathrm{OC}$$

where OC is termed the ***sagitta*** of the arc QOP, and OC is given by

$$\mathrm{OC} \, . \, \mathrm{CD} = \mathrm{CP}^2.$$

But CD $\doteqdot 2R$, so that the above formula ensues at once.

[It must be remembered that the formula for PN is only true if the curved surface actually touches the flat one: in practice, this may be prevented by dust, and the surfaces are distorted if each exerts a force upon the other. Experimentally, we shall see how this difficulty is overcome, by using a difference method.]

Let us calculate the time for a light disturbance to travel from X, Fig. 29·10 (*c*)—a point in the lens—to P and back to X; also via P to N and back again to X. It must be remembered that when light is reflected at a medium-air interface there is no change

in phase introduced: when the reflexion occurs at an air-medium interface a phase retardation equivalent to a path retardation of $\frac{\lambda}{2}$ is introduced. In the present instance this latter type of reflexion occurs at N and the time for the disturbance to travel via N is increased by $\frac{1}{2}$T, where T is the period of vibration. In effect, the path via N has been increased by $\frac{1}{2}\lambda$.

If c_0 is the velocity in air, c the velocity in the material of

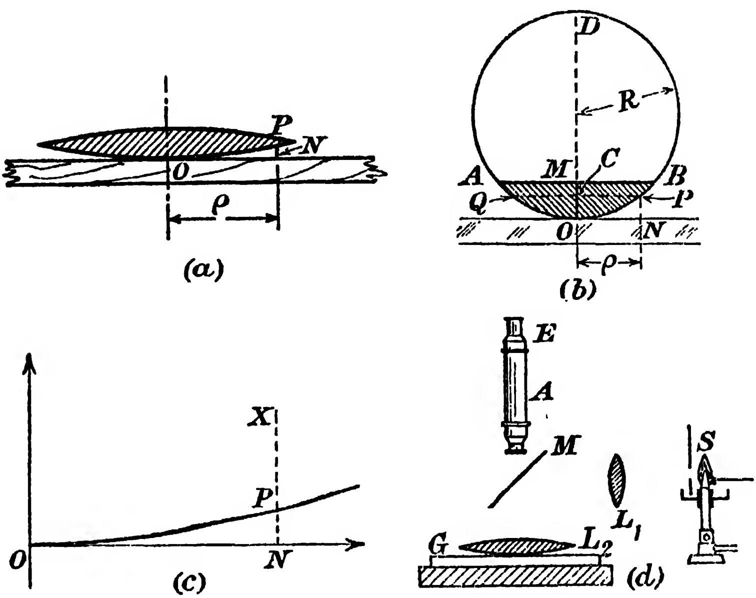


FIG. 29·10.—Newton's Rings.

the lens, then the time of transit from X and back to X by a path wholly within the lens is

$$\frac{XP + PX}{c}.$$

For the path XPNPX, it is

$$\frac{2 . PN}{c_0} + \frac{T}{2} + \frac{XP + PX}{c}.$$

The time difference is therefore

$$\frac{2PN}{c_0} + \frac{T}{2}.$$

Now in this time the light would travel in air a distance

$$\left(2\frac{\mathrm{PN}}{c_0} + \frac{\mathrm{T}}{2}\right)c_0 = 2\mathrm{PN} + \frac{\lambda}{2},$$

where λ is the wave-length in air, i.e. in the medium of the film. For darkness, this equivalent path difference must be equal to $(m + \frac{1}{2})\lambda$, where m is a positive integer including zero, i.e.

$$2\mathrm{PN} = m\lambda \quad \text{or} \quad \frac{\rho_m^2}{\mathrm{R}} = m\lambda.$$

The centre of the system $(m = 0)$ is therefore black.

If ρ_1 is the radius of the smallest dark ring (not the centre itself)
ρ_m " " " mth " "

$$\frac{\rho_m^2 - \rho_1^2}{\mathrm{R}} = (m - 1)\lambda$$

$$\therefore \lambda = \frac{\rho_m^2 - \rho_1^2}{(m-1)\mathrm{R}} = \frac{\delta_m^2 - \delta_1^2}{4(m-1)\mathrm{R}},$$

where δ_m is the diameter of the mth dark ring.

Similarly, if ρ_n is the radius of the nth dark ring, we have

$$\lambda = \frac{\rho_m^2 - \rho_n^2}{(m-n)\mathrm{R}}$$

For brightness, the above time difference must be $m\mathrm{T}$, where m is an integer not including zero, i.e.

$$\frac{2\mathrm{PN}}{c_0} = (2m-1)\frac{\mathrm{T}}{2}, \quad \text{or} \quad \frac{\rho_m^2}{\mathrm{R}} = (2m-1)\frac{\lambda}{2}.$$

Newton's Rings by Transmitted Light.—When the rings are viewed by transmitted light, the system of fringes is complementary to that formed by reflected light. It must be remembered, however, that in the rings formed by transmitted light, the intensity of the ray passing through without reflexion is much greater than that which suffers two reflexions—in fact, its intensity is almost equal to that of the incident light, i.e. the amplitudes of the interfering wave trains are not equal and the minima are therefore not completely dark.

The above means that the transmitted rings are much more indistinct than the rings obtained by reflected light. In the formation of these last-mentioned rings, there is interference between two rays each of which has undergone reflexion at one of the boundaries of the film : in consequence, they have the same amplitude and the minima are completely dark.

The experimental arrangement for viewing Newton's rings is indicated in Fig. 29·10 (d). A circular opening, S, in a metal screen

is illuminated with sodium light [1] and a converging lens L_1 is placed so that S is at its focus. This position of the lens may be found with the aid of a plane mirror as in the experimental determination of the focal length of a converging lens. The parallel beam of light produced by L_1 is reflected from a thin microscope cover glass M on to the lens L_2. A microscope, A, previously adjusted so that the upper surface of the glass plate G is in focus [or rather some small scratch on the surface is in focus], is moved until the rings are seen. With the arrangement here described it is a little difficult to find the rings. They may be found more readily by using an extended flame as the source, i.e. remove S and L_1. This source should be as near as is convenient to M. The rings are very brilliant under these conditions and the glass M is adjusted until an eye E vertically above the point of contact of L_2 and G sees an image of the flame. The microscope is, of course, removed during this procedure. Under these conditions the rings will also be seen and the microscope may be brought into position. If the rings are measured under these circumstances the error due to the fact that the light is not parallel is, in general, negligible, since the rings occupy only a small area around the point of contact so that the rays producing them may be regarded as being almost parallel. If it is essential to measure the rings using light which is strictly parallel the aperture S and lens L_1 may be introduced after the rings have been found. They will still be in the field of view, only very much reduced in brilliancy.

To avoid locating the position of the centre of the system of rings, and some of the difficulties mentioned above, it is better to measure the diameters of the rings. If δ_m and δ_n are the diameters of the *m*th and *n*th bright rings respectively,

$$\begin{aligned}\delta_m^2 - \delta_n^2 &= 4\left[(2m-1)\frac{\lambda R}{2} - (2n-1)\frac{\lambda R}{2}\right] \\ &= 4(m-n)\lambda R.\end{aligned}$$

Deduce the value of λ from observations on the 10th and 5th, 9th and 4th rings, etc.

In the above treatment we have neglected the fact that the rings which are formed in the air film are examined after refraction through the lens L_2, but if this lens is thin the error thus introduced is very small. The error is avoided completely by placing the flat plate on top of the curved surface and viewing the rings by reflected light.

[1] This is readily obtained by heating a mixture of common salt and borax on the end of a gas-mantle support. A lid from a small tin is fitted over the top of the bunsen burner to prevent any of the molten mixture from falling on the bench—cf. sketch.

Alternative Treatment for a Study of Newton's Rings.—Let us suppose that a flat plate rests on top of a spherical surface with a large radius of curvature, R, so that an air film of variable thickness is formed as indicated in Fig. 29·11 (*a*). Consider light falling normally on the film and that the light reflected from the bounding surfaces of the film is considered. On account of the phase change equivalent to an increase in path of $\frac{1}{2}\lambda_0$, where λ_0 is the wave-length of the light in the film (air), there will be darkness at the point of contact, D_c, between the surfaces. This will be surrounded by the first bright ring at B_1, and B_1A_1, the thickness of the film, will be such that

$$2B_1A_1 = \tfrac{1}{2}\lambda_0.$$

This is because in passing from a dark fringe to a bright one in an

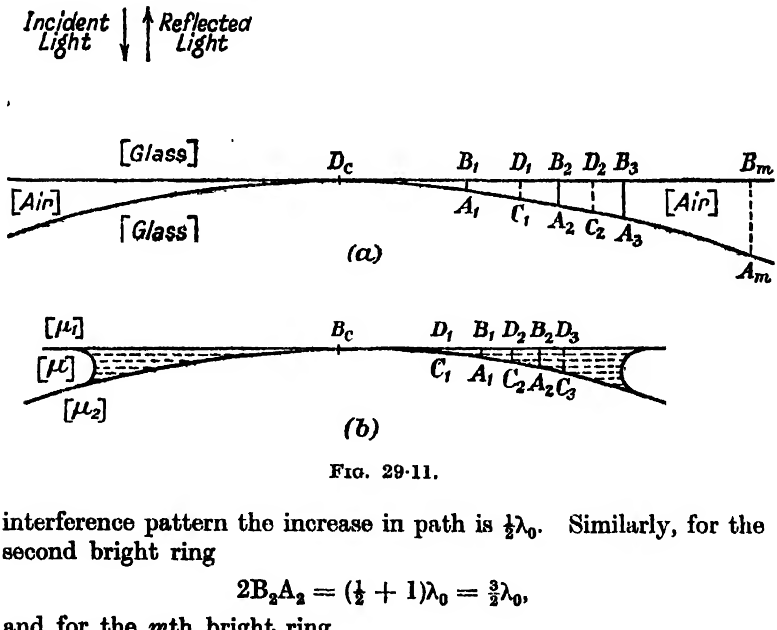


FIG. 29·11.

interference pattern the increase in path is $\frac{1}{2}\lambda_0$. Similarly, for the second bright ring

$$2B_2A_2 = (\tfrac{1}{2} + 1)\lambda_0 = \tfrac{3}{2}\lambda_0,$$

and for the *m*th bright ring

$$2B_mA_m = (2m - 1)\frac{\lambda_0}{2}.$$

But if ρ_m is the radius of the *m*th bright ring $\rho_m{}^2 = B_mA_m . 2R$, so that

$$\rho_m{}^2 = \tfrac{1}{2}R\lambda_0(2m - 1) = R\lambda_0 m - \frac{R\lambda_0}{2}.$$

If we call $\rho_m{}^2 = y$, $m = x$, and plot a series of observations, we ought to obtain a straight line whose slope is $R\lambda_0$. It is of course

better to measure the diameters of the bright rings since the centre of the system cannot be located with accuracy, when the above equation will require modification.

Let us now suppose that a liquid of refractive index μ is trapped between the above surfaces and let us assume that μ_1, the refractive index of the material of the plate, and μ_2, that of the material of the lens, are such that $\mu_1 > \mu > \mu_2$. The phase change is zero on reflexion at each of the boundaries so that there is a bright spot at the centre of the system. Proceeding as before we have :—

For the first bright ring $2B_1A_1 = \lambda$.
" second " " $2B_2A_2 = 2\lambda$.
. . . . .
" *m*th " " $2B_mA_m = m\lambda$.

Hence

$$\rho_m{}^2 = 2R.\frac{m\lambda}{2} = \frac{mR\lambda_0}{\mu}.$$

If $\rho_m{}^2 = y$ and $m = x$, the slope of the straight line is $\frac{R\lambda_0}{\mu}$. Thus the ratio of the slopes of the two graphs gives a value for the refractive index of the liquid comprising the film.

[N.B.—If $\mu_1 < \mu < \mu_2$, a phase change equivalent to an increase in path of $\frac{1}{2}\lambda$, where λ is the wave-length of the light in the medium of refractive index μ, takes place at both the upper and lower reflecting surfaces : thus a bright spot is the centre of the ring system.]

Determination of the Angle between Two nearly Parallel Plates.—Let OA and OB, Fig. 29·12, be the faces of two parallel

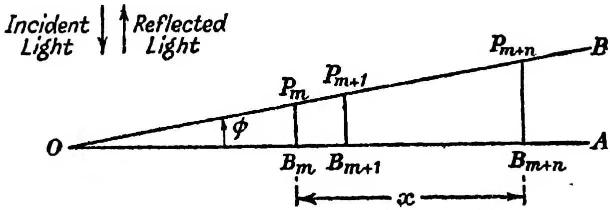


FIG. 29·12

plates of glass enclosing an air film and let ϕ be the very small angle between these faces. When the plates are viewed by reflected monochromatic light a system of equidistant interference fringes will be observed : each fringe will be parallel to the line of intersection of the two faces. Suppose a bright fringe occurs at B_m : the next bright fringe will occur at B_{m+1}, where

$$P_{m+1}B_{m+1} \quad P_mB_m = \tfrac{1}{2}\lambda,$$

and the thickness will increase by $\frac{1}{2}\lambda$ for each successive bright fringe. Hence, if there are n bright fringes between B_m and B_{m+n}, we have

$$\phi = \frac{P_{m+n}B_{m+n} - P_mB_m}{B_{m+n}B_n} = \frac{n\lambda}{2x}$$

where ϕ will be given in circular measure.

EXAMPLES XXIX

1.—Interference bands are produced by a Fresnel's biprism in the focal plane of a microscope. This plane is 100 cm. from the slit. A converging lens placed between the biprism and the microscope gives two distinct images of the slit in two positions. These images are 2·86 and 4·13 mm. apart in the two instances. If the mean distance between the interference fringes is 0·1002 cm., what is the wave-length of the light used?

2.—Explain the production of the fringes which may be seen in a soap film formed on a vertical wire frame. If the fringes seen with transmitted sodium light are 5 mm. apart, what is the angle between the surfaces of the film?

[$\mu = 1{\cdot}33$, $\lambda_D = 5{\cdot}89 \times 10^{-5}$ cm.].—(L. '25.)

3.—Describe in detail a method of producing interference fringes and show how it may be applied to determine the wave-length of the light transmitted by a piece of red glass.

4.—A slightly convex lens is placed on a plane glass plate. Account for the formation of the rings seen round the point of contact when it is looked at normally. How would you use such an arrangement to prove that the wave-length of red light is greater than that of blue light?

5.—Describe and give the theory of an accurate method of determining the wave-length of sodium light.

6.—A convex surface of known radius of curvature is placed in contact with a plane surface. Derive an expression for the radius of a ring as seen by reflected light in terms of the angle between the two reflecting surfaces at the points in question.

7.—In an experiment with Newton's rings, using reflected light, the diameters of two consecutive rings are 2 cm. and 2·02 cm.; what is the radius of curvature of the lens surface in contact with the plane glass? [λ for light used = 5897A.]

8.—A plane diffraction grating has a constant $1{\cdot}79 \times 10^{-4}$ cm. First and second order spectra are found when the angles of diffraction are 18° 40′ and 39° 48′ respectively when monochromatic light is incident normally on the grating. Calculate the wave-length of the light.

9.—Show that if monochromatic light of wave-length λ is incident on a plane diffraction grating at an angle i, the angular positions of the lines in the diffraction pattern are given by the equation

$$m\lambda = (a + b)(\sin i + \sin \theta),$$

where the symbols have their usual meanings.

If $i = 45°$, calculate values for the angular positions of the different-order spectra for light of wave-length $6{\cdot}0 \times 10^{-5}$ cm., if the grating constant, i.e. $(a + b)$, is $(6 \times 10^3)^{-1}$ cm.

10.—Monochromatic light of wave-length $5{\cdot}46 \times 10^{-5}$ cm. falls normally on a grating consisting of parallel wires equidistant from one another. Spectra are observed in positions 17′ and 52′ from the zero-order spectrum. Calculate values for (*a*) the grating constant, (*b*) the thickness of the wire.

11.—A drop of liquid of refractive index 1·58 is placed between a plano-convex lens and a plane glass plate on which the former rests. The materials of the plate and lens are identical and have a refractive index 1·50. Calculate the diameter of the fifth bright ring, if the radius of curvature of the lens surface is 200 cm. and light of wave-length $5{\cdot}9 \times 10^{-5}$ cm. is reflected normally from the system. What effect will be observed when the plate is replaced by another whose material has a refractive index 1·65 ?

12.—The refracting angles of a biprism are each 0·025 radian ; the material of the biprism has a refractive index 1·59 ; the slit and screen are 10 cm. and 80 cm. respectively from the biprism. Calculate a value for the separation of the fringes produced by light of wave-length 5900 A.

13.—The convex surface (radius of curvature 400 cm.) of a plano-convex lens rests on a concave surface of 500 cm. radius and Newton's rings are viewed by reflected light of wave-length $5{\cdot}9 \times 10^{-5}$ cm. Calculate the diameter of the fifth bright ring seen.

14.—Calculate a value for the angular separation of the two sodium lines ($\lambda_1 = 5896$ A, $\lambda_2 = 5890$ A) in the first- and second-order spectra produced by a plane diffraction grating having 5000 lines per centimetre. The light is incident normally on the grating.

15.—Parallel light from a mercury-vapour lamp falls normally on a plane diffraction grating having 5000 lines per centimetre. The diffracted light is focused on a screen by a converging lens of focal length 100 cm., the axis of the lens being orientated so that the mercury green line in the first-order spectrum is at the second principal focus of the lens. Calculate values for the angular positions in the first-order spectrum of the lines corresponding to light of the following wave-lengths: 5791 A, 5770 A, 5461 A (green), 4358 A, 4339 A.

Also calculate their distances, on the screen, from the green line.

CHAPTER XXX

POLARIZED LIGHT

Introductory.—In developing the wave theory of light in the previous chapters no stipulation as to the nature of the waves has been made. Moreover, it has always been assumed that all light behaves in the same way under the same circumstances. For example, if a surface reflects light, it will always reflect it. It will now be shown, however, that there are certain rays of light possessing a two-sidedness so that a water or other surface may refuse to reflect them at a certain angle of incidence. Such light is said to be ***polarized.*** The only type of polarized light we shall discuss is that in which the vibrations occur in parallel planes—it is termed ***plane polarized*** light. The study of such light is important because it helps us to decide whether or not light consists of longitudinal or transverse vibrations in the æther.

Let us first consider an analogy. If a wire spring about 6 ft. in length [such as is used for hanging curtains] is stretched slightly and passed through a narrow slit the following experiment may be performed :—When the spring is plucked by displacing it parallel to the length of the slit the disturbance set up passes through the slit. If it is plucked in a direction at right angles to the slit the disturbance fails to pass through it. On the other hand, a longitudinal disturbance excited in the spring always passes through the slit irrespective of the orientation of the slit with respect to the spring.

Returning to the optical problem, a section of a tourmaline crystal of the shape shown in Fig. 30·1 (*a*) is obtained. Such a section is said to be cut parallel to the crystallographic axis AB. All crystals possess certain axes to which their shape and other properties may be referred in the simplest possible manner. These axes are termed the crystallographic axes, but the student must be careful not to regard them as fixed positions in a crystal. They are merely directions, and any line parallel to a crystallographic axis is an identically equivalent axis in all respects. If a sheet of white paper is viewed through such a tourmaline slice more than 1 mm. thick the emergent light will be tinged green owing to selective absorption in the crystal, but the intensity of the transmitted light

will not be affected when the crystal is rotated about an axis normal to the face of the crystal receiving the light. If a second similar crystal is held so that the crystallographic axes of the two crystals are parallel, no appreciable difference in the light transmitted by them will be noticed. On rotating the second [or first] crystal as previously the intensity of this light will gradually diminish until when the axes of the crystals are mutually perpendicular no light is transmitted. The two crystals are said to be ***crossed.*** Fig. 30·1 (*b*) illustrates this complete extinction of light by crossed tourmalines. This experiment may be explained if we assume that the light transmitted by such a slice of a tourmaline crystal has become plane polarized, for such light can only be transmitted by the second crystal when their axes are similarly orientated. From the analogy given above it is clear that light waves must be transverse and not longitudinal.

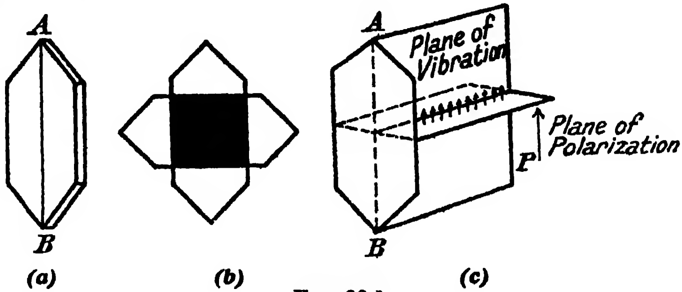


FIG. 30·1.

Now although we have proved that light consists of a transverse motion in the æther and that plane polarized light is obtained when these vibrations are in one plane we do not know definitely the plane in which they occur. Some writers have assumed that the vibrations are parallel to the crystallographic axis, and others that they are in a direction normal to this. The electromagnetic theory of light teaches us that something is probably going on in each of these planes, but for our present purpose we shall assume that in the plane polarized light transmitted through a tourmaline crystal the vibrations are in a plane containing the crystallographic axis. A plane at right angles to this is called the ***plane of polarization***—cf. Fig. 30·1 (*c*).

The above peculiar property of tourmaline is due to the fact that when a ray of light is incident normally on the face of the crystal it divides into two rays, one polarized in the plane containing the axis AB and the other in a plane perpendicular to the above. The former of these is absorbed while the latter is permitted to pass, i.e. light in which the vibrations are parallel

to the crystallographic axis AB emerges from the tourmaline plate.

Polarization by Reflexion.—M and N, Fig. 30·2, are two pieces of plate glass which can be rotated about horizontal axes, the lower one N, in addition, being capable of rotation about a vertical axis. Light from an arc lamp passes through a small aperture S placed at the principal focus of a converging lens L so that a parallel beam of light falls on M. The direction of this beam of light and the tilt of M are arranged so that the angle of incidence is 55°, and that the reflected beam passes vertically downwards on to N. This second mirror is parallel to M so that it receives light at an angle of incidence equal to 55°. A circular patch of light will be seen on a cylindrical translucent screen placed round N. When N is rotated round a vertical axis the intensity of the reflected light falls gradually to zero when the rotation has amounted to 90°. On continuing the rotation in the same direction the intensity will become a maximum when the rotation is 180°, fall to zero at 270°, and finally be a maximum again when a complete rotation has been made. If N is rotated rapidly a ring of light with two maxima will be produced on the screen. This can be explained if the light reflected from a glass surface at an angle of 55° is plane polarized.

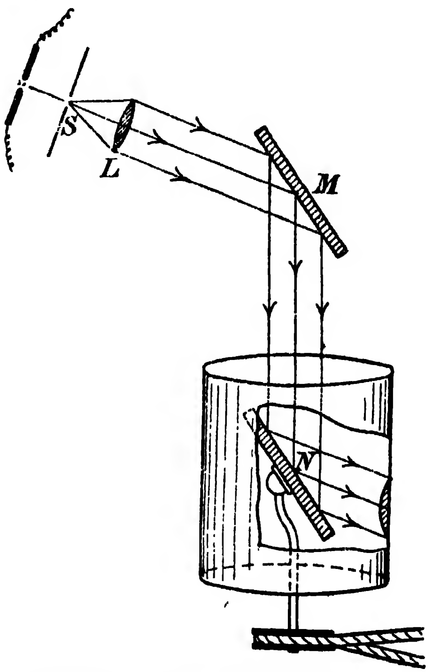


Fig. 30·2.—Polarization by Reflexion.

Brewster's Law.—When light is incident at an angle other than 55° the light reflected from the lower mirror is never reduced to zero although there are variations in intensity. From a series of experiments made with different reflectors, Brewster discovered that complete extinction occurred when the reflected ray was normal to the refracted ray. The particular angle of the incidence when extinction is possible is termed the ***angle of polarization*** or ***polarizing angle.*** If OA, OB, and OC, Fig. 30·3, are the incident

reflected, and refracted rays when the angle of incidence, a, is equal to the angle of polarization, Brewster's law states that

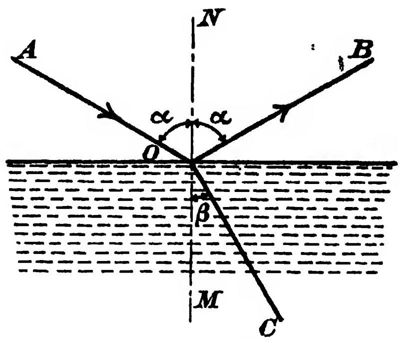


FIG. 30·3.—Brewster's Law.

$$\widehat{BOC} = \frac{\pi}{2}, \text{ i.e. } (a + \beta) = \frac{\pi}{2}.$$

But

$$\mu = \frac{\sin a}{\sin \beta} = \frac{\sin a}{\cos a} = \tan a.$$

Since μ depends upon the colour of the light it follows that complete polarization in the reflected ray is only obtained with monochromatic light.

Double Refraction.—ERASMUS BARTHOLINUS, a native of Denmark, first drew attention to some remarkable properties of Iceland spar or calcite ($CaCO_3$). The same phenomena may be observed in all crystals except those belonging to the cubic system. A crystal of calcite is bounded by six rhombohedral faces termed the cleavage faces of the crystal. They are so called because they represent directions in the crystal along which it can most easily be split. If such a rhomb is placed over a small blot of ink on white paper two images will be seen. Crystals of calcite are said to exhibit ***double refraction.*** HUYGHENS made a careful study of these rays and found that one of them obeyed Snell's law of refraction. He termed this the ***ordinary ray***; the other did not obey this law and he termed it the ***extraordinary ray.***

Let S, Fig. 30·4, be a small luminous object lying on the normal SN to a rhomb of Iceland spar (calcite). Let A_1SA_2 be a small pencil of rays from S. At A_1 the ray SA_1 gives rise to ***two*** refracted rays, A_1B_1 and A_1X_1. These are the ordinary and extraordinary rays respectively. At B_1 and X_1 these are refracted along B_1C_1 and X_1Y_1, both emergent rays being parallel to SA_1, the incident ray. We see, therefore, that the refractive index for the extraordinary ray is less than that for the ordinary ray.

When we consider the ray SA_2, however, we find that there is formed by refraction the ordinary ray A_2B_2, which, at B_2, emerges as B_2C_2. The extraordinary ray is A_2X_2 and it must be noted that even if SA_1 and SA_2 are equally inclined to ON, the two extraordinary rays ***in*** the rhomb are ***not*** equally inclined to SN. (In fact, they are not, in general, in the plane of the diagram, but, for the purpose of the explanation here given, this point is ignored.) The emergent ray X_2Y_2 is, however, parallel to SA_2. We therefore have two divergent pencils emerging from the rhomb. If these

enter an eye situated on SN produced, two images will be seen, (*a*), the ordinary image O, lying on the normal SN, (*b*), the extraordinary image E, ***not*** on the normal SN. Hence, when the rhomb is rotated about SN, the image O remains stationary, while E rotates.

If a microscope is focused in turn on O and then on E, it will be found that O is nearer to the rhomb than is E.

Moreover, if parallel light is incident upon the rhomb, only one image is seen.

If, in the case of divergent rays, the emergent light is examined

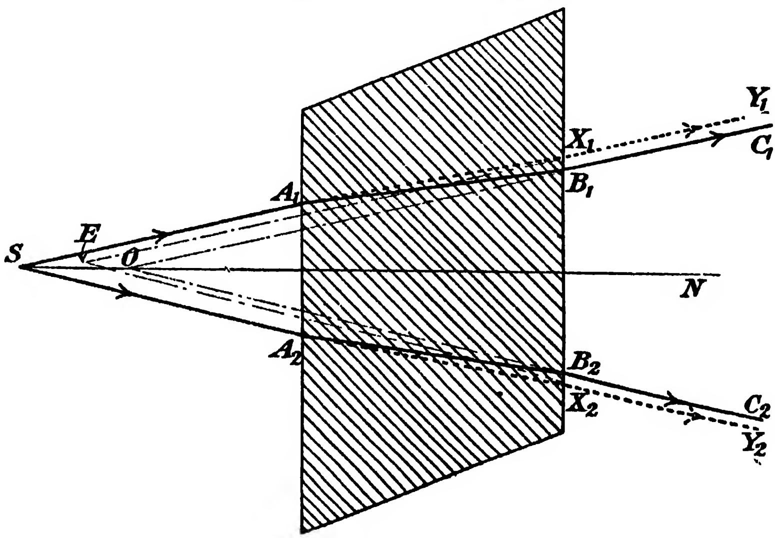


FIG. 30·4.

after passing through a tourmaline crystal, on rotating the tourmaline about a horizontal axis it will be found that in some positions only the ordinary ray is transmitted, while in a position at right angles to the above only the extraordinary ray gets through the crystal. This experiment proves that the ordinary and extraordinary rays are vibrating in planes mutually at right angles.

The Nicol Prism.—A very convenient means of obtaining plane polarized light or of detecting it is the so-called Nicol prism: a better name for it would be a Nicol rhomb. It is constructed from calcite in the following manner:—A calcite crystal is illustrated in Fig. 30·5 (*a*). At the point A the angles of the three faces are equal and obtuse. The straight line OA drawn through A and

equally inclined to these three faces is termed the ***optic axis.*** Now since in crystals all parallel lines are strictly equivalent as far as the properties of the crystal are concerned, any line parallel to the one we have drawn is an optic axis. An optic axis of a

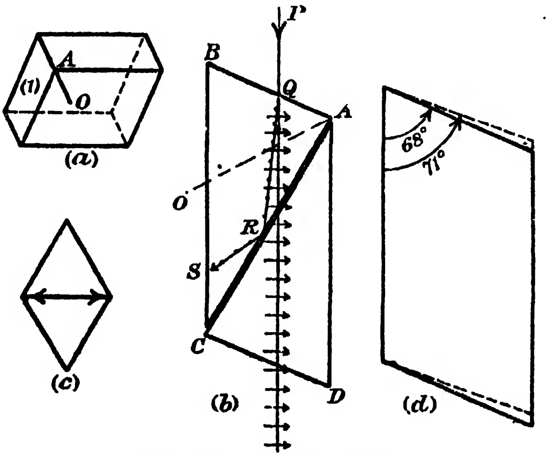


FIG. 30·5.—A Nicol Prism.

crystal possesses the property that plane waves of light whose normals are parallel to the optic axis travel through the crystal with a velocity which is independent of the direction of vibration of the light, i.e. the extraordinary and the ordinary wave fronts remain together, or the corresponding rays are coincident. Any plane normal to the face of a crystal and containing an optic axis is called a ***principal plane*** for that face of the crystal. [The case is strictly analogous to that of a 'normal to a plane surface' which is a direction and not a particular line.]

Let ABCD, Fig. 30·5 (*b*), be the principal plane for light incident on the face AB. In constructing a Nicol rhomb it is found necessary for AB not to be a natural face which makes an angle of 71° with BC but to be an artificial face formed by grinding away the natural face until the $\widehat{ABC}$ is 68°. The rhomb so formed is then cut in two by a plane at right angles to the principal plane—its trace is AC. The two halves are afterwards cemented together by Canada balsam. When a ray PQ falls on the face AB the ordinary ray suffers total internal reflexion at R along RS since the refractive index of the balsam [1·55] is less than that of calcite for the ordinary ray [1·658]. On the other hand, the extraordinary ray is transmitted when the prism is constructed in the above manner. The emergent light is plane polarized, the vibrations being parallel to AB, the shorter diagonal of the face receiving the incident light.

The faces of the prism parallel to BC or AD are blackened to absorb stray radiations.

A section of the calcite before and after grinding is shown in (*d*). The length of a nicol is about three times its width so that a minimum amount of material is used in its manufacture.

The Production and Analysis of Polarized Light.—The most convenient method whereby plane polarized light may be obtained is with the aid of a Nicol rhomb. The nicol is mounted in front of a source of light when the transmitted light is plane polarized, the vibrations being parallel to the shorter diagonal of the face receiving the light—cf. Fig. 30·5 (*c*). This is equivalent to stating that the vibrations are parallel to the principal plane of the prism or that the plane of polarization of the emergent light is perpendicular to the principal plane, i.e. the plane of polarization is normal to the plane of the paper in Fig. 30·5 (*b*). When a nicol is used in this way it is termed a ***polarizer.***

To detect plane polarized light a second nicol is placed to receive the light under examination. If this nicol has its principal plane parallel to the plane in which the vibrations are taking place in the incident light, no apparent change will be seen on looking through this prism. But if the nicol is rotated through 90° the field will be dark, since the light vibrations will be in such a plane that the light cannot pass through. If no change takes place the light is not plane polarized. We cannot say, however, that the light is unpolarized, for it is possible for light to be polarized in another way and yet pass through the nicol. Such light is said to be ***circularly*** polarized and other means have to be developed to distinguish circularly polarized and unpolarized light. When a Nicol is used to examine light in this manner it is termed an ***analyser.***

Rotation of the Plane of Polarization.—Let two nicols, N_1 and N_2, Fig. 30·6, be arranged with their principal planes parallel. Parallel light from a source S is then transmitted through the

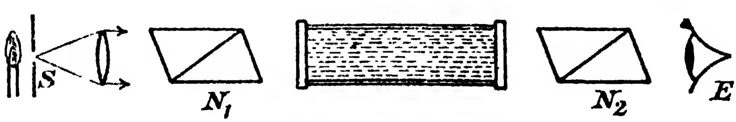


FIG. 30·6.—Principle of a Simple Polarimeter.

system. [It is necessary to use parallel light so that the ordinary rays produced in the first section of the nicol shall all fall on the Canada balsam at an angle of incidence greater than the critical angle.] If N_2, the analyser, is rotated through 90° the field appears dark and the nicols are said to be ***crossed.*** If, however, a tube

containing a substance such as turpentine is introduced between the crossed nicols the field will, in general, appear bright. To establish darkness it is necessary to rotate the analyser by an amount depending on the length of the tube and the particular substance in it.

Quartz, aqueous solutions of cane sugar, ice, etc., are capable of rotating plane polarized light: they are said to be ***optically active***. The rotation is said to be ***right-handed*** if, on looking along the path of the light towards its source, the rotation appears to be clockwise. If it is anti-clockwise it is termed a ***left-handed*** rotation. To decide whether or not a substance exhibits a right- or left-handed rotation experiments must be made with two tubes of different lengths: a calculation of the rotation per cm. will indicate the correct sign, for if the correct one has been chosen the rotation will be directly proportional to the length.

The ***specific rotation*** is defined as the angle of rotation in degrees per unit length of the optically active column, divided by the mass of substance per unit volume. Thus for D light and at temperature t, for pure a liquid

$$\left[\alpha\right]_{D}^{t} = \frac{\theta}{l\rho},$$

where θ is the angle of rotation, in degrees, produced by a column of length l, and ρ is the density of the liquid.

Now consider a solution containing m gm. of active substance in $(m + M)$ gm. of a solution. Then there will be m gm. of active material in $(m + M)\rho^{-1}$cm.^{3} of solution, where ρ is the density of the solution.

$$\left[\alpha\right]_{D}^{t} = \frac{\theta}{l} \div \frac{m}{(m + M)\rho^{-1}} = \frac{\theta}{lc},$$

where c is the concentration in gm.cm.$^{-3}$

[N.B.—l is often expressed in decimetres.]

The Influence of Strain on Polarized Light.—If a piece of glass which has been carefully annealed is placed between crossed nicols the field remains dark. If, however, the piece of glass is strained by bending, the field will no longer be dark. This experiment indicates a means of detecting strain in glass. In some of the cheaper forms of glass blocks the presence of strain may be strikingly shown in this way.

EXAMPLES XXX

1.—When sodium light is incident on the surface of a plate of block glass, the reflected light is plane polarized when the angle of incidence is 58° 23′. What is the refractive index of the glass?

2.—Discuss the evidence for the belief that light vibrations are transverse. What is plane polarized light? Describe two methods of obtaining it.

PART IV

ACOUSTICS

CHAPTER XXXI

WAVE MOTION AND THE NATURE OF SOUND

Introduction.—The term ***sound*** is used to denote the sensation we receive by means of our ears and also to name the physical cause of this sensation. In physics we have to deal with the external disturbance and our present aim is to investigate the manner in which it arises, the mode in which it travels to us, and the cause of the differences which enable us to distinguish one sound from another. Everyday experiences teach us that the source of the sound is in a state of vibration so that the preliminary part of our work must be a study of vibrating bodies. We have already learnt that a particle executes a simple harmonic motion when it is displaced from its zero or rest position if forces directly proportional to the displacement arise tending to restore it to its zero position.

Graphical Representation of a S.H.M.—It has already been shown [cf. p. 36] that a S.H.M. may be regarded as the projection of a uniform circular motion on any diameter of the circle. Thus, if

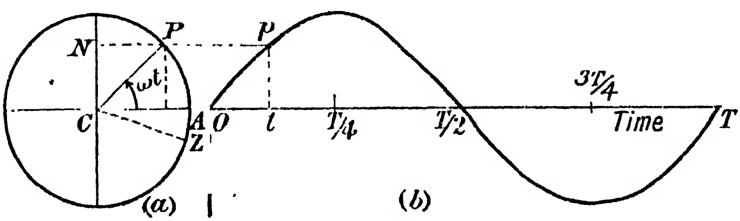


FIG. 31·1.—Graphical Representation of a S.H.M.

P, Fig. 31·1 (a), is moving with uniform angular velocity, ω, in a circle whose centre is C, and PN is drawn perpendicular to the y-axis, the point N performs a S.H.M. To represent this motion graphically O, Fig. 31·1 (b), is taken as origin, the time being represented along OT, while the displacement of N from the centre C at any particular instant is given by the corresponding ordinate. Thus if at time t

the rotating particle has moved to a point P, where ACP $= \omega t$, the position of its projection N on the y-axis is represented by p, where O$t = t$, and $pt =$ NC. If a series of points corresponding to the positions of N at various instants during one complete period of its oscillation are determined in this way and joined by a smooth curve the graph shown in (b) is obtained. Such a curve is called a ***harmonic curve.*** Since $pt =$ NC $=$ CP sin $\omega t = a$ sin ωt, where a is the radius of the circle, it follows that the equation to the curve we have obtained is

$$y = a \sin \omega t.$$

It is now customary to write this equation as

$$y = \hat{a} \sin \omega t,$$

where $\hat{a} = a$, and is the ***amplitude*** or ***peak value of the displacement.***

If we had considered the motion of the projection of P on the x-axis we should have had

$$x = \hat{a} \cos \omega t.$$

If, instead of measuring the time from one particular instant when the point P crosses the x-axis, it is measured from an instant when P is at Z, where $\widehat{ZCA} = \phi$, then $\widehat{ZCP} = \omega t$, i.e. $\widehat{ACP} = \omega t - \phi$, and the equation to the curve is

$$y = \hat{a} \sin (\omega t - \phi)$$

The angle $(\omega t - \phi)$ is called the ***phase*** of the vibration, while $-\phi$ is termed the ***initial phase*** or ***epoch.***

Similarly

$$x = \hat{a} \cos (\omega t - \phi).$$

Resultant of Two S.H.Ms. in the same Straight Line.—Let us suppose that the point N, Fig. 31·1 (a), continues to move with S.H.M. about the point C as centre while the book itself moves parallel to the y-axis with a S.H.M. of the same period but different amplitude and phase. The resultant of these two motions gives the actual displacement of N with reference to the table on which the book may be resting. The thin curve in Fig. 31·2 represents the displacement of N, its equation being $y = \hat{a} \sin \omega t$. If the motion of the book alone is $y = \hat{b} \sin \left(\omega t + \frac{\pi}{2}\right)$, i.e. its phase is exactly a

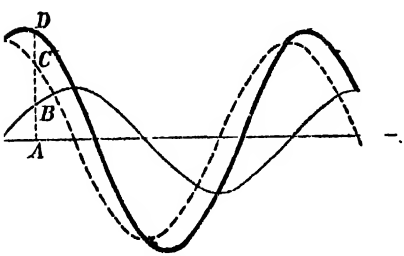


FIG. 31·2.—Composition of S.H.Ms. in the same straight line. (Same period.)

quarter of a period in advance of the first, then the dotted line represents this motion. At the instant represented by A the displacement of N due to the first motion is AB and AC due to the second. The total displacement is therefore AD where AD = AB + AC. Similarly, at any other instant the resultant displacement is the algebraic sum of the ordinates of the two curves. The thick curve represents the motion of N due to the combined action of the two S.H.Ms. The figure shows that the resultant of two such motions of equal period and in the same straight line is also a S.H.M., and this particular instance may be generalized analytically as follows :—

Resultant displacement $= \hat{a} \sin \omega t + \hat{b} \sin (\omega t - \phi) = y$ (say), where ϕ is the difference in phase between the two superimposed motions. Then

$$\therefore y = \hat{a} \sin \omega t + \hat{b} \sin \omega t \cos \phi + \hat{b} \cos \omega t \sin \phi$$
$$= \sin \omega t(\hat{a} + \hat{b} \cos \phi) + \cos \omega t(\hat{b} \sin \phi)$$
$$= \hat{c} \sin(\omega t + \theta), \text{ if } (\hat{a} + \hat{b} \cos \phi) = \hat{c} \cos \theta \text{ and } \hat{b} \sin \phi = \hat{c} \sin \theta$$

$$\text{or, } \tan \theta = \frac{\hat{b} \sin \phi}{\hat{a} + \hat{b} \cos \phi}, \text{ and } \hat{c}^2 = \hat{a}^2 + \hat{b}^2 + 2\hat{a}\hat{b} \cos \phi.$$

The same artifice can be employed when the S.H.Ms. have different periods. Fig. 31·3 has been constructed when the ratio of the two periods is 5 : 4 ; the dotted curve represents the vibration of shorter period and it has been assumed that at the commencement of the observations the two motions are producing their maximum displacements in the same direction. The resultant vibration is no

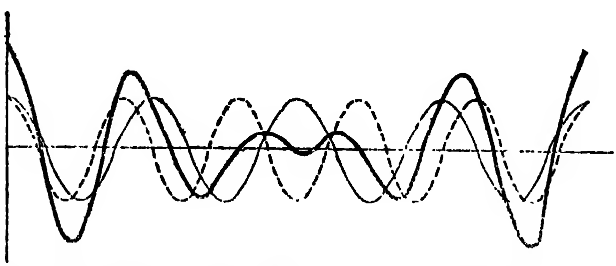


FIG. 31·3.—Composition of S.H.Ms. in the same straight line. (Different periods.)

longer simple harmonic, for its amplitude alternates between large and small values. The resultant amplitude is very large at the beginning when the two motions are in the same phase, while at the time when the faster has executed 2·5 complete vibrations and the slower 2 the two motions are out of phase, i.e. the difference in phase is π, the resultant amplitude is small. After an equal lapse of time the two are again in phase and the re-

sultant amplitude is a maximum. It is clear from the diagram that the displacement is a maximum whenever the slower particle has executed four complete vibrations, or the faster one five.

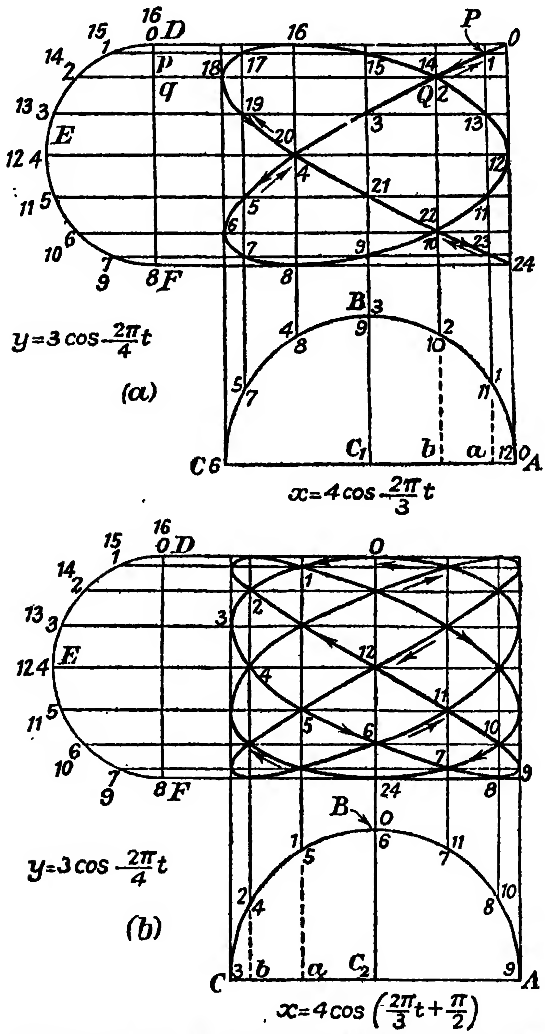


FIG. 31·4.—Composition of S.H.Ms. at Right Angles.

Resultant of Two S.H.Ms. at Right Angles.—Let us suppose that two S.H.Ms. at right angles to one another are simultaneously impressed upon a particle, e.g. if these two directions are chosen as reference axes, let $x = 4\cos\frac{2\pi}{3}t$ and $y = 3\cos\frac{2\pi}{4}t$ be the equations

to the component motions. These indicate that the amplitudes are as 4 : 3, whilst their periodic times are as 3 : 4. Two semicircles ABC and DEF, Fig. 31·4 (a), with their radii in the ratio of the amplitudes, are drawn. Let their arcs be divided into six and eight equal parts, the dividing marks being numbered in each instance consecutively and in an anticlockwise direction. This is done since the tracing points for each motion will describe each of these portions in the same time. [When the x- and y-tracing points are actually on those portions of the auxiliary circles not shown, the corresponding numbers are placed inside the circles.] Since the motions are initially in phase the tracing points must occur at the extreme ends of their swings in the positive directions at the same instant. Let us suppose that time is measured from the occasion when one such particular event occurs. The particle will therefore start at O. At the end of the interval corresponding to the positions of the tracing points marked 1, the displacement in the x-direction will be Aa, whilst that in the y-direction will be Dp ; the particle will therefore be at P [also marked 1]. At the end of the second interval the particle will be at Q, etc. Proceeding in this way the complete curve representing the motion of the particle is obtained. The arrows on the curve denote the direction in which the tracing point moves during the first half of the time taken for the figure to be described once. The arrows at the side indicate the direction of motion during the second half of the above time.

If the two impressed motions are represented by $x = 4\cos\left(\frac{2\pi}{3}t + \frac{\pi}{2}\right)$ and $y = 3\cos\frac{2\pi}{4}t$, the amplitudes and periodic times are in the same ratio as above, but now the x-motion is initially in phase $\pi/2$ ahead of the y-motion, i.e. when the tracing point giving the y-motion is at D, the other is at B—Fig. 31·4 (b). Suitable dividing marks are again marked on the auxiliary circles, and numbered consecutively in an anticlockwise direction as before. This is done in order to ascertain the direction of the motion of the point upon which the two motions are impressed. [If the x-component had been given by $x = 4\cos\left(\frac{2\pi}{3}t + \frac{3\pi}{2}\right)$, the resulting figure would have been identical with that just obtained, but the sense of description would be opposite—hence the importance of adopting the correct mode of numbering the marks on the auxiliary circles.]

The curves in Fig. 31·5 have been obtained in the same way : they are more simple than those of Fig. 31·4, since the amplitudes of the components are equal and the ratio of the periods is either (1) unity or (2) two.

Blackburn's Pendulum.—A thin string about 7 ft. long has its two ends fixed to a horizontal rod E, Fig. 31·6. This string is cut at its centre and the two ends attached to a heavy lead ring, B,

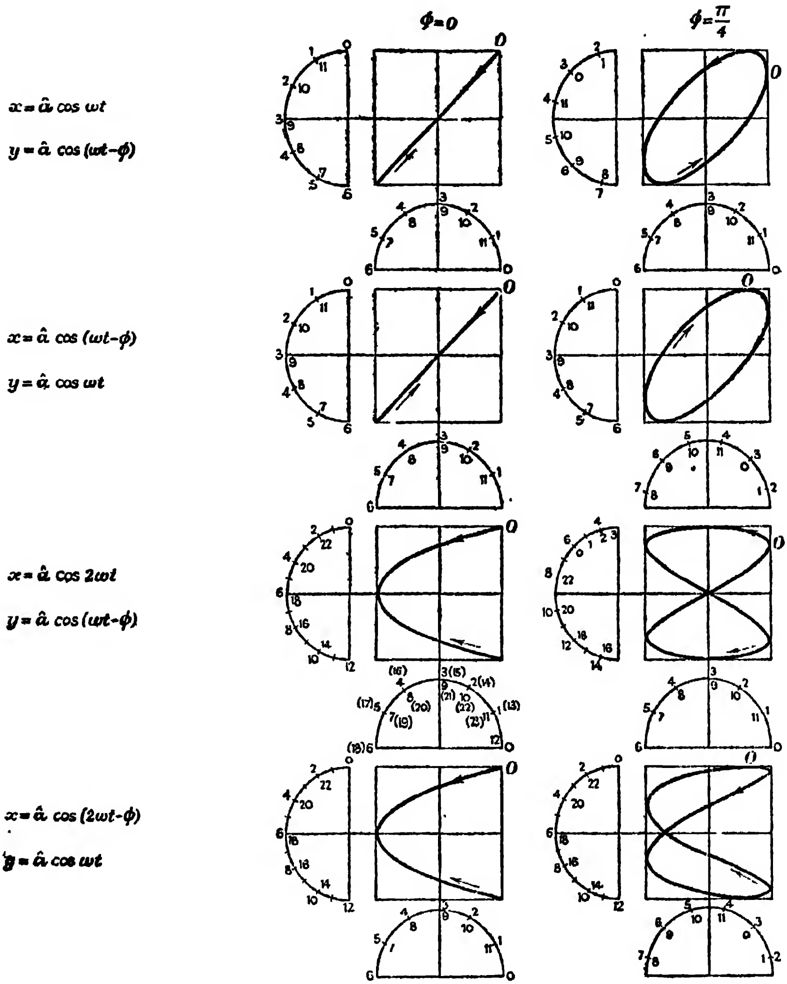
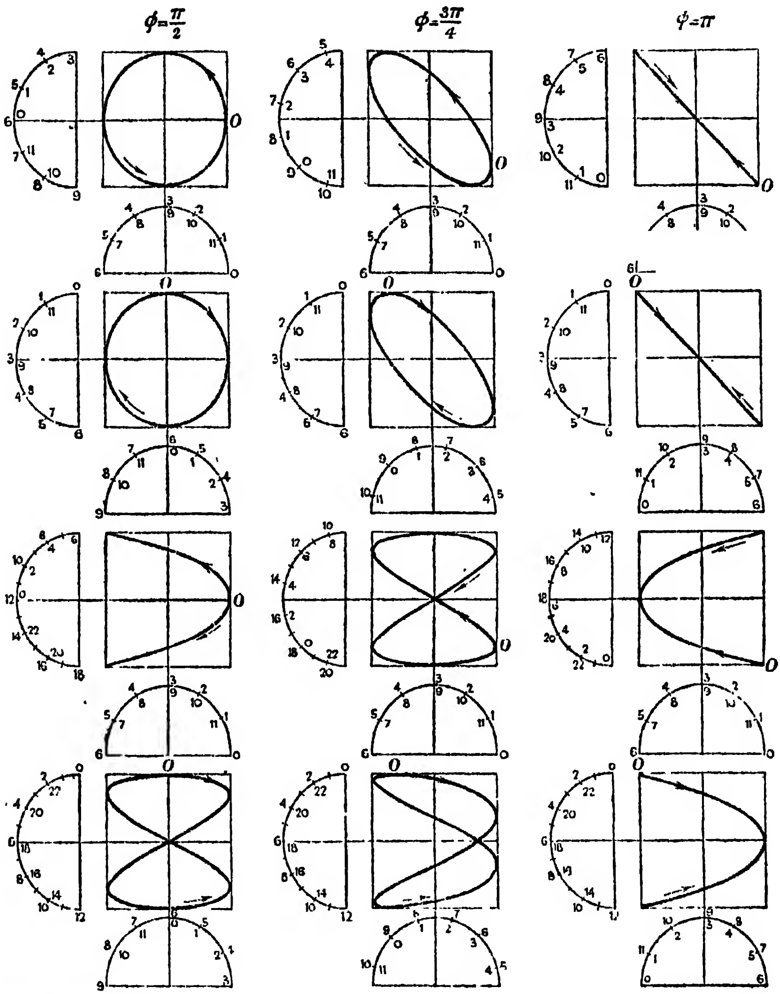


FIG. 31·5.-

carrying a glass funnel whose exit tube has been constricted at its lower end. A clip A enables the string to be caught up as shown. The whole forms a pendulum of length EB for vibrations perpendicular to the plane of the figure, but one of length AB when the vibrations are in the plane of the figure. The periods may be adjusted by altering the position of A. When the bob

is displaced outwards in a slanting position and then released, its motion is that of two S.H.Ms. mutually perpendicular. If some dry sand is placed in the funnel and permitted to escape



Lissajous' Figures.

on to a sheet of paper immediately below, a record of the motion may be obtained. The lead ring, having a mass considerably in excess of that of the sand, keeps the centre of gravity of the system constant so that its period does not vary appreciably.

Lissajous' Figures.—Two metal strips having different lengths or differing in nature are mounted so that one, A, may

oscillate in a horizontal plane and the other, B, in a vertical plane. An aluminium screen with a small hole drilled in it is attached to the first strip and a convex lens to the other. The distance apart of the screen and lens is adjusted so that a clear image of the illuminated aperture is formed on a screen. If A alone vibrates the image is drawn out into a horizontal line, while if B oscillates by itself the image becomes a vertical line. Both strips perform S.H.Ms. so that when both oscillate together a curved figure more complicated than those in Figs. 31·4 and 31·5 is generally obtained. Such curves are known as *Lissajous'* figures. They enable us to compare the frequencies of two vibrating objects. For example, let A and B be the prongs of two tuning-forks one of which has a frequency n while the other (B) has a frequency slightly different from this. At some instant the difference in phase will be $\frac{\pi}{2}$ when a closed curve will be observed. Since the ratio of the frequencies of the two forks differs slightly from unity this pattern will not persist and various curves will appear in turn until the slower fork has made one complete oscillation less than the other, when the phase difference will again be $\frac{\pi}{2}$. Let t be the time which elapses between two successive appearances of the same closed curve of light. During this time one fork has made ft vibrations while the other has made $ft \pm 1$. To determine the correct sign to be used in this equation a small piece of wax is attached to the fork B and the time elapsing between successive reappearances of the same curve of light noted. If this is less than before it follows that the difference in frequency between the two forks has been increased so that the frequency of B is less than that of A. The frequency of the fork B is therefore

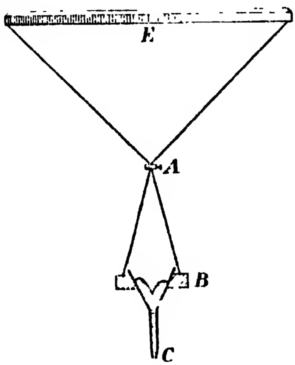


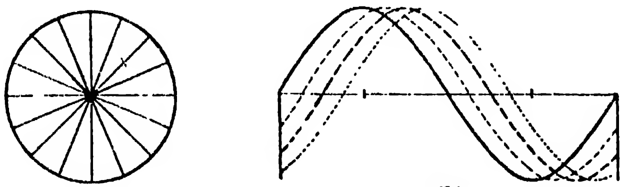
FIG. 31·6.—Blackburn's Pendulum.

$$\left(\frac{ft-1}{t}\right) = \left(f - \frac{1}{t}\right).$$

Experiment.—Use Lissajous' figures to ascertain when the periods of two strips are identical. Measure the length, λ_1, of A (say), and then alter its length to λ_2 until the frequency is doubled. Verify that

$$\lambda_1 = \sqrt{2}\,\lambda_2.$$

Waves.—Hitherto we have only considered the vibration of an isolated particle or point in a body. In Nature examples are often found of whole bodies each of whose particles is vibrating with S.H.M. but in which the phase differs regularly from one particle to the next. As an illustration let us consider a row of particles lying in a straight line and equidistant from one another when they are at rest. Suppose that these particles all execute a S.H.M. of common amplitude but in which the successive phases differ by $\frac{2\pi}{16}$ or 22·5°. Then the motions of all these particles may be represented by those of sixteen particles moving at equal distances round a circle the radius of which equals the common amplitude of the vibrations, for the displacement of any particle at a particular instant is given by the ordinate of the corresponding particle in the circle—cf. Fig. 31·7 (*a*). The positions of the particles when the phase of the first has increased by successive multiples of $\frac{2\pi}{16}$ are



(*a*) (*b*)

FIG. 31·7.

indicated in Fig. 31·7 (*b*). It will be noticed that all the curves may be obtained from the first by displacing this to the right so that, as the particles in the circle continue to rotate, the curve representing their positions moves to the right. These curves are similar in shape to the harmonic curves already discussed [cf. p. 579] but differ in at least one important respect. The harmonic curve represents the position of ***one*** particle at different times in its history, whereas the curves now contemplated are really instantaneous pictures representing the positions of ***all*** the disturbed particles at one particular instant; they are termed ***wave-form*** curves.

The effect produced as the wave-form curve moves forward is known as a ***wave-motion,*** a very familiar example of which occurs whenever a stone is dropped into the still water of a pond. The stone depresses the water at the point of contact: but since water is an elastic medium, forces arise tending to restore the surface of the water to its original level. The magnitude of the depression is therefore finite. When the water regains its normal level for the first time it possesses considerable inertia in consequence of which it 'overshoots the mark' and produces an elevation. The restoring forces soon bring the moving water to rest momentarily,

after which the inertia increases, the water passes beyond its original position and a depression is formed: this process is repeated. [In actual practice the amplitude is gradually reduced by viscous forces and the disturbance soon dies away.] But meanwhile, however, the disturbance is not confined to the point where the stone entered the water. As each particle of water is displaced its neighbour is influenced and begins to participate in this up-and-down motion. The resultant effect is that although the particles move up and down a wave passes across the surface of the water. When the displacement of a particle is a maximum in the positive direction, i.e. above the normal level of the water, that particle is said to be at the ***crest*** of a wave, whilst, when it is a maximum in the other direction, it is at a ***trough.*** The distance from one crest (or trough) to the next crest (or trough) is termed a ***wave-length.*** Particles whose abscissæ differ by one wave-length (or an integral number of wave-lengths) are in the same phase. [The phases of two vibrations are the same when they are identical or differ by an even number of π radians.]

From the above remarks we see that a wave-motion may be regarded as a disturbance which travels in a medium and which arises from parts of the medium executing definite periodic vibrations about their mean positions.

The ***velocity*** with which the disturbance is propagated is the distance through which it moves in unit time: the ***frequency*** is the number of complete waves passing a fixed point per unit time. If the wave-length is λ, and the frequency f, the velocity c is expressed by

$$c = f\lambda.$$

Transverse Waves.—Let a heavy 'bob' hang from a light helical spring and let AX, Fig. 31·8 (a), be a long cord of negligible mass attached to it. It will be assumed that the string rests in a horizontal position and distances measured from A along the string will be denoted by x. When the bob is displaced slightly from its position of rest it will execute a S.H.M. of amplitude $\hat{a}$ and period T, say. The effects of damping are neglected. Then the motion of the end A of the string may be obtained from the uniform circular motion of a point P_0, the radius of the circle being $\hat{a}$ and the angular velocity ω, where $\omega = \frac{2\pi}{T}$. If time is measured from the instant the bob leaves its zero position the portion of the cord disturbed after t seconds will be ABC, Fig. 31·8 (b), where P_0 has moved to P to make $P_0\widehat{O}P = \omega t$, so that the displacement of A is given by

$$y = \hat{a} \sin \omega t.$$

After the lapse of a longer time the cord will be displaced as indicated in Fig. 31·8 (c), and a wave-length, λ, is the distance indicated, i.e. λ is the distance through which the disturbance advances when the reference point P_0 makes one complete revolution.

To obtain the equation to the motion of any point B, at distance x from A when at rest, we note that the time required for the disturbance to travel a distance x is $T.\frac{x}{\lambda}$, since λ is the distance traversed in time T. Hence the initial position of the tracing point Q in the reference circle which eventually gives the motion of B, cf. Fig. 31·8 (d) must be such that $Q_0\widehat{O}P_0 = \omega . T\frac{x}{\lambda} = 2\frac{\pi}{\lambda}.x = \eta x$, say. Hence the motion of B is represented by the equation

$$y = \hat{a}\sin(\omega t - \eta x) = \hat{a}\sin(\omega t - \phi),$$

if $\phi = \eta x$, $-\phi$ is known as the initial phase of the motion.

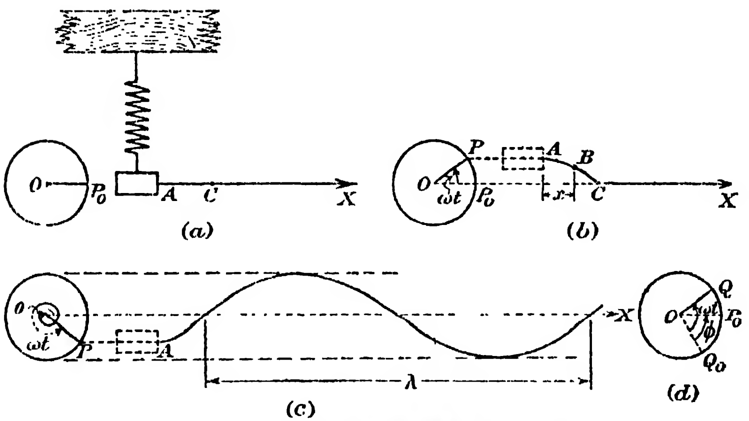


Fig. 31·8.—Transverse Waves along a Cord.

It must be noted that the particles in the cord move up and down, but the disturbance is propagated to the right. Each particle in the string executes the same motion as the moving 'bob,' but as we have shown, not necessarily synchronous with it, for the phase of any given particle in the string depends on its distance from A and the velocity with which the disturbance travels forward. Such a progressive disturbance is termed a ***transverse wave*** since the vibrating particles move in a direction at right angles to that along which the disturbance travels. The important point to be emphasized is that the particles vibrate about a mean position but their kinetic energy travels forward.

If we measure time from some other zero, this equation becomes

$$y = \hat{a} \cos (\omega t - \eta x),$$

and in advanced work it is preferable to use this as the wave equation. It is still better to write

$$\xi = \hat{a} \cos (\omega t - \eta x),$$

where ξ is the displacement which may not be perpendicular to x; the use of y for displacement frequently leads to the supposition that the displacement must be perpendicular to the x-axis.

Longitudinal Waves.—Let us now assume that a piston, P, nearly fitting a tube, as shown in Fig. 31·9, executes a S.H.M. parallel to the length of the tube. The air at P is alternately compressed and rarefied as the piston moves towards and away from it. When it is compressed the particles tend to relieve the strain which has been created by compressing the adjacent layers of air to the right. These in turn hand on the compression. A few moments later the strip moves to the left and a rarefaction takes place at P, a condition which is passed on to the adjacent layers as before. Owing to the definite time

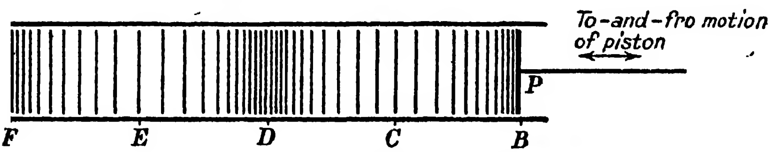


FIG. 31·9.—Longitudinal Waves.

interval between these two states and the compressions and rarefactions immediately following them a wave-motion consisting of these alternate compressions and rarefactions is propagated forward. Such a wave is said to be a ***longitudinal wave*** since the vibrating particles move in a direction parallel to that along which the wave advances. Again we have to notice that, as with transverse waves, the vibrating particles only move through a small distance about a mean position but the energy is carried forward.

The Wave Equation.—When a particle is executing a simple harmonic motion, that motion may be represented analytically by the equation

$$\begin{aligned} \xi &= \hat{a} \cos \omega t \\ &= \hat{a} \cos 2\pi f t, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)} \end{aligned}$$

where ω is the angular velocity of the point moving in a circle of radius a defining the motion; ω is termed the ***angular or circular frequency,*** while $f = \frac{1}{T}$ is the frequency of the motion.

Each of the above expressions involves the assumption that the motion begins at the time $t = 0$ and the displacement is a maximum

when $t = 0$. An equation of a more general nature, but still representing a simple harmonic motion, is

$$\xi = \hat{a}\cos(\omega t - \phi) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

From this, $\xi = \hat{a}$ first when $\omega t - \phi = 0$, i.e. $t = \frac{\phi}{\omega}$. In other words the motion represented by (ii) is reckoned from a time $\frac{\phi}{\omega}$ later than that in (i).

Suppose the equation $\xi = \hat{a}\cos\omega t$ represents the motion of a point A, on a cord, along which waves are travelling with velocity c—a cord is used merely for the sake of a concrete example. Now the disturbance at A reaches a point B at distance x from A, at a time $\frac{x}{c}$ after it was at A. Suppose that when the displacement at A is given by $\xi = \hat{a}\cos\omega t$, that at B is given by $\xi = \hat{a}\cos(\omega t - \phi)$. Then at a time $\frac{x}{c}$ later the displacement at B is the same as that at A at time t, i.e.

$$\underset{\left\{t=t+\frac{x}{v}\right\}}{\xi_B} = \hat{a}\cos\left[\omega\left(t + \frac{x}{c}\right) - \phi\right] = \underset{t=t}{\xi_A} = \hat{a}\cos\omega t.$$

i.e. $$\frac{\omega}{c}x = \phi.$$

But $\omega = \frac{2\pi}{T}$ and $c = \frac{\lambda}{T}$, so that the above condition becomes

$$\frac{2\pi}{\lambda}x = \phi \text{ or } \eta x = \phi, \text{ where } \eta = \frac{2\pi}{\lambda}.$$

The equation $\xi = \hat{a}\cos(\omega t - \eta x)$ therefore gives the displacement, in general, at any point on the cord. It contains two variables, t and x. The equation representing the motion of a given particle is obtained by considering x as constant, i.e.

$$\xi = \hat{a}\cos(\omega t - \kappa_1) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iv)}$$

On the other hand, the displacement of all points on the cord is, at a particular instant, given by

$$\xi = \hat{a}\cos(\kappa_2 - \eta x) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(v)}$$

From equations (iv) and (v) it is clear that a ***cinematographic picture*** of a selected point is identical with that of an ***instantaneous picture*** of the whole wave (cord).

The equation $\xi = \hat{a}\cos(\omega t - \eta x)$ represents a wave disturbance travelling along the x-axis in the positive direction. If the wave travels in the opposite direction the equation becomes

$$\xi = \hat{a}(\cos\omega t + \eta x)$$

because the disturbance is at B before it reaches A, i.e. the disturbance at B at time $t - \frac{x}{c}$ is identical with that at A at time t.

For longitudinal waves we still write

$$\xi = \hat{a}\cos(\omega t - \eta x),$$

but it must be noted that ξ is parallel to the x-axis.

Velocity of Longitudinal Waves.—NEWTON first proved that the velocity of longitudinal waves, in a medium whose modulus of

elasticity for the particular type of strain set up is E and whose density is ρ, is expressed by $\sqrt{\frac{E}{\rho}}$. In rods, the compressions and rarefactions cause temporary alterations in length so that Young's modulus for the particular material concerned determines the magnitude of the elastic forces which arise. Hence, for these,

$$V = \sqrt{\frac{E}{\rho}}.$$

In liquids and gases changes in volume are produced by the compression and it is owing to the volume elasticity that the waves are propagated. It has already been shown [cf. p. 144] that the volume elasticity for an ideal gas under isothermal conditions is equal to the pressure p which it exerts, so that $c = \sqrt{\frac{p}{\rho}}$. This result was obtained by Newton. Since one atmosphere equals $76 \times 13{\cdot}6 \times 981$ dyne. cm.$^{-2}$ and the density of air at 0° C. is 0·001293 gm.cm.$^{-3}$, we obtain by substitution $c = 2{\cdot}80 \times 10^4$ cm.sec.$^{-1}$. Experimental determinations of the velocity of sound in air show that it is $3{\cdot}32 \times 10^4$ cm.sec.$^{-1}$. The difference between these values is too great to be attributed to experimental errors, but the problem remained an enigma to Newton. About a century later LAPLACE detected the cause of the discrepancy. He realized that the compressions and rarefactions in a gas take place so rapidly that the gas is considerably heated during a compression and cooled during a rarefaction. It is therefore not the isothermal elasticity but the adiabatic elasticity which controls the process. The adiabatic equation is $pv^\gamma =$ constant, where γ is the ratio of the specific heat of the gas at constant pressure to that at constant volume. Let v be the volume of gas at pressure p, and suppose that when the pressure becomes $(p + \delta p)$, the volume is $(v + \delta v)$, the change taking place adiabatically.

Then $pv^\gamma = (p + \delta p)(v + \delta v)^\gamma$,

i.e.
$$1 = \left(1 + \frac{\delta p}{p}\right)\left(1 + \frac{\delta v}{v}\right)^\gamma,$$

$$= \left(1 + \frac{\delta p}{p}\right)\left(1 + \gamma.\frac{\delta v}{v}\right), \left[\begin{array}{l}\text{since } \delta v^2 \text{ and higher powers} \\ \text{of } \delta v \text{ are negligible}\end{array}\right],$$

$$= 1 + \frac{\delta p}{p} + \gamma.\frac{\delta v}{v}., \qquad \left[\text{since } \delta p.\delta v \text{ is negligible}\right].$$

$$\frac{\delta p}{p} + \gamma.\frac{\delta v}{v} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$\text{Now elasticity} = \frac{\text{increase in stress}}{\text{strain}} = \lim_{\delta p \to 0} \frac{\delta p}{-\left(\frac{\delta v}{v}\right)} = -v.\frac{dp}{dv}.$$

But from (1), under the adiabatic conditions contemplated,

$$\lim_{\delta p \to 0} \left(-v.\frac{\delta p}{\delta v}\right) = \gamma p.$$

$$\therefore \qquad \text{Adiabatic elasticity} = \gamma p.$$

For air, $\gamma = 1{\cdot}40$, so that, at 0°C.,

$$c = \sqrt{\frac{\gamma p}{\rho}} = \sqrt{1{\cdot}40} \times 2{\cdot}80 \times 10^4 = 3{\cdot}31 \times 10^4 \text{cm.sec.}^{-1}.$$

Since this agrees with the velocity of sound as determined experimentally we conclude that during the passage of longitudinal waves in air the variations in pressure take place under adiabatic conditions.

The Velocity of Sound in an Ideal Gas and its Variation with Pressure and Temperature.—The expression $c = \sqrt{\gamma p/\rho}$ may be written $c = \sqrt{\gamma p v} = \sqrt{\gamma \mathscr{R} T}$, if v is the specific volume of the gas, T its temperature on the absolute scale, and $\mathscr{R}$ is the gas constant. Since γ and $\mathscr{R}$ are constants, the above formula shows that the velocity of sound in an ideal gas is independent of its pressure (for non-explosive sounds) but directly proportional to the square root of its absolute temperature. Experiment shows this result holds for gases such as air, hydrogen, oxygen, etc.

Velocity of Sound in Free Air.—The velocity of sound in free air has been the subject of many experimental investigations during the last three or four centuries. The first determination which can be considered at all reliable was made about the middle of the eighteenth century by three members of the French Academy. Two stations about 30 kilometres apart were selected and at constant intervals of time during the night cannons were fired, one at each station. An observer at the other station determined the time elapsing between seeing the flash from the explosion and hearing the report. The distance between the stations having been measured accurately, the velocity of the sound was deduced. Their results indicated that the velocity of sound in air increased with increase in temperature but was independent of the pressure. They also showed that the velocity was independent of the actual distance of the observer from the source but that it was increased when it travelled with the wind and diminished when it travelled against it. By taking the mean time of propagation in two opposite directions as in these experiments the wind effect was eliminated.

The important objection to be urged against all such determinations as the above is that three errors arise which are very difficult to eliminate—together they comprise what is termed the 'personal equation.' Three distinct factors arise in making a determination of the time interval: the flash is *seen*, the sound is *heard*, and a chronometer is *operated* by the observer's finger. For the first two of these time is required for the brain to interpret the signal received, whilst before the last operation can be carried out a message must be sent from the brain to the observer's finger. The time interval or lag peculiar to any person (or instrument) between the recording of an event and its perception is known as the '*personal equation*' for that person or instrument.

The error due to the personal equation of the observer may be very much reduced by employing mechanical means to record the arrivals of the various signals, but this does not eliminate the error entirely, for all pieces of such apparatus have their own 'personal equations'; it is, however, much more constant than that of an observer and may be evaluated and then eliminated by making experiments over widely different distances. REGNAULT attempted an important series of experiments on these lines in 1864.

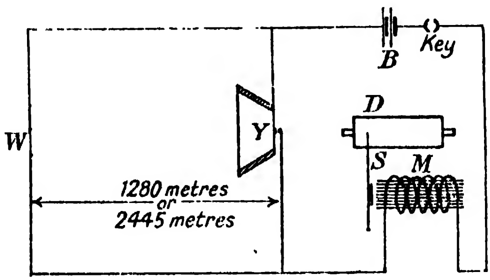


FIG. 31·10.—Regnault's Apparatus for Measuring the Velocity of Sound in Air.

The reciprocal firing of guns at stations first 2,445 metres and then 1,280 metres apart was the method employed. The apparatus shown in Fig. 31·10 was duplicated. D was the drum of a chronograph revolving at a constant and known speed. S was a style making a trace on the drum. A gun fired at the first station broke the wire, W, forming part of the electrical circuit shown: the style moved to the left leaving on the drum a definite indication of the time of origin of the sound waves. The sound-waves on arrival at the other station were incident upon a wooden cone over the end of which a membrane was stretched. This membrane moved when the sound-waves arrived and temporarily completed the electrical circuit, so that the electromagnet was again excited,

and the end of S made a mark on the drum perpendicular to the general trace. Since the speed of the drum was known, the time interval for the sound to travel across WY could be computed.

By determining the velocity of sound waves in the opposite direction and calculating the mean of the velocities in the two directions a value for the velocity of sound in air independent of the effect of any wind was obtained. REGNAULT found, however, that this apparatus had a personal equation comparable with that of a trained observer: it was eliminated by making experiments over the two different distances already mentioned.

STONE, in 1871, at Cape Town where he was Astronomer Royal, attempted to eliminate the error due to the personal equation in the following way :—Two observers were stationed 641 ft. and 15,499 ft. respectively from a gun which was fired. Each recorder reported the arrival of the sound at his station on an electric chronograph situated at the observatory. The difference between the times of the arrival of the sound as recorded by each observer was the time required for the sound to travel across the distance separating them, but slightly in error owing to the fact that their 'personal equations' were not likely to be the same, especially as the intensity of the sound perceived by each was different. To eliminate the difference between these two personal equations, a smaller gun was fired at such distances from the two observers that the intensity of the sound was approximately the same as in the main experiment. The distances were now 162 ft. and 1,483 ft. from the gun. The time for sound waves to travel across the distance between the two observers was then calculated from the provisional value for the velocity of sound deduced from the first experiment when no correction was applied. It was found to be 1·177 sec. The recorded time interval was 1·265 sec., i.e. it was in excess of the computed time by 0·09 sec. This represents the difference between the two personal equations, i.e. the correction for the difference between the personal equations is —0·09 sec. When this correction was applied, Stone obtained

$$c_0 = 332{\cdot}4 \text{ metre. sec.}^{-1} = 1{,}090{\cdot}6 \text{ ft. sec.}^{-1}.$$

GREELY, working in the Arctic regions where conditions are sometimes very still and low temperatures cause the water content of the air to be small, found that the velocity of sound in air could be represented by the equation $c = (332 + 0{\cdot}6t)$ metre. sec.$^{-1}$ where t is the temperature on the centigrade scale.

Accurate Determination of the Velocity of Sound in Air.—The velocity of sound in air may be determined by methods which are classified as direct or indirect. In the direct method, the time taken for a sound to travel across a measured distance is determined,

and we have to consider the relative advantages of using a long or a short distance. The main objections to the long distance method are:

(i) Very intense sounds have to be used—e.g. the discharge of a cannon, and it is doubtful whether the velocity near the source is the same as that some distance away.

(ii) It is impossible to apply corrections for wind, temperature, and humidity with any great degree of accuracy.

(iii) The 'personal equation' of an observer or of a recording device is involved.

In the short distance method (i) and (ii) are avoided, while (iii) depends on the method adopted. In REGNAULT's experiments a gun was used. To free the experiment from the personal equation of the observer, both the discharge of the gun at one station and the arrival of the sound at the other were recorded electrically. There was nothing to guarantee, however, that the recording device did not possess a 'personal equation' of its own. Simultaneous firing from both stations was employed to eliminate the effect of the wind.

Hebb's Telephone Method for Measuring the Velocity of Sound in Air.—HEBB, in 1905, at the suggestion of MICHELSON, devised the following method for measuring the velocity of sound. A_1 and A_2, Fig. 31·11, were two paraboloidal mirrors arranged coaxially. Waves of sound were sent out from a source at the focus F_1 of A_1, and collected at the focus F_2 of A_2. T_1 is a telephone transmitter near F_1 and T_2 a second telephone transmitter at F_2. Each transmitter is in series with a battery B_1 or B_2, respectively, and one of

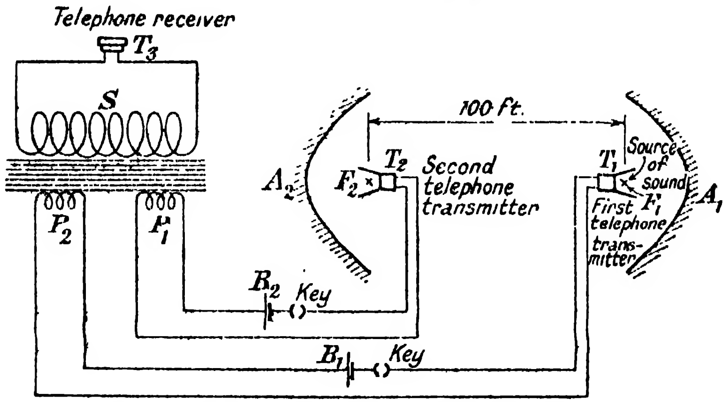


FIG. 31·11.—Hebb's Telephone Method for Measuring the Velocity of Sound in Air.

the primaries P_1 or P_2 of a special induction coil having two primaries. The secondary of this coil was connected to a telephone receiver.

Now suppose that waves are sent out by a source at F_1. Some pass directly to T_1 and set its diaphragm in vibration with a definite phase relation depending on the distance of T_1 from F_1. Other waves are collected at F_2 and operate the transmitter T_2—the phase of its vibrations will depend on the distance $F_1A_1A_2F_2$. The vector sum of the two effects will be given in the receiver T_3. If we assume that the intensities (amplitudes) of these effect are equal, it is possible, by

moving one mirror parallel to itself, to change the relative phases of the two effects so that they will alternately annul and reinforce one another. This affords a good method of measuring the wave-length of the sound. If the mirror can be moved through a distance of 100 wave-lengths, and each position of maximum resultant effect in T_3 determined with an error not exceeding 0·1 of its wave-length, the wave-length may be determined correctly to within one part in a thousand.

The mirrors were 5 feet in diameter and had a focal length of 15 inches. They were made of plaster of Paris. The source of sound was a tube 0·75 in. in diameter, closed at one end, and having a stream of air blown across the other. It was arranged so that as few overtones as possible were present. The pitch was adjusted to be equal to that of a standard fork and could be maintained constant to within 1 part in 5,000. The hall in which the experiments were carried out was 120 feet long, 10 feet wide and 14 feet high. There was no wind and the mean temperature was deduced from the indications of six thermometers arranged alternately on the walls. The final result obtained for the velocity of sound in air at 0° C. was

$$c_0 = 331{\cdot}46 \text{ metre. sec.}^{-1}.$$

The Velocity of Sound in Water.—In 1826, COLLADON and STURM measured the speed of sound waves travelling through the water of Lake Geneva. A bell was supported in the water from a boat and sound waves excited by striking the bell with a lever: the moment the bell was struck the same lever fired a charge of gunpowder. Since the experiments were carried out at night the flash from the explosion was seen by a distant observer who held a form of ear-trumpet in the water. The end in the water was closed by a membrane which vibrated when the sound-waves reached it. The upper end of the trumpet was placed against the observer's ear so that the arrival of the sound was easily detected. From the interval of time between the flash and the arrival of the sound, and the distance over which it had travelled, the velocity of the sound was calculated.

The Velocity of Sound in Sea-Water.—If the velocity of sound in any medium is known, the distance between two points in it can be determined, if the time required for sound to travel from one point to the other and through the medium is known—in fact, the problem is the reverse of the one with which we are now dealing. Although there are more accurate means of surveying on land, the method has great possibilities when used for surveying at sea. It is fairly easy to determine the position of a ship at sea with respect to two land stations, but it is much more difficult for a ship to ascertain its position by the ordinary methods of trigonometrical surveying, especially in rough weather owing to the rolling and pitching of the boat. Moreover, this method fails utterly in foggy weather when it is most essential that the captain of the ship should know his whereabouts. The Radio-Acoustic method was developed by the British Admiralty for this purpose. But before describing it let us see how the velocity of sound in sea water was measured.

A wireless operator on board one of two ships stationed at a known distance apart used a double key to fire an electrical detonator placed in a submerged charge of gun-cotton, and at the same time to transmit a wireless signal. An observer on the second ship received the wireless signal—the transmission of which across a short distance may be considered instantaneous—while some seconds later the sound was received. From the known distance over which the sound had travelled and the time taken the velocity of sound in sea-water became known. At any temperature t° C. it is given by $c = (4{,}756 + 14t)$ ft. sec.$^{-1}$.

With this knowledge at hand a ship may locate its position in cloudy or foggy weather by emitting simultaneously two signals: the one a sound signal through the sea, and the other a wireless signal. These are then picked up by two land stations at a carefully surveyed distance apart. Suppose that T is the interval of time between the reception of the two signals at one land station. If D is the distance of the ship from the station, and c_0 and c the velocities of wireless waves and sound in sea water, respectively,

$$\mathrm{T} = \frac{\mathrm{D}}{c} - \frac{\mathrm{D}}{c_0}.$$

Now c_0 is equal to the velocity of light (3×10^8 metre. sec.$^{-1}$); its reciprocal, when compared with the reciprocal of c, is zero for all practical purposes and we may write

$$\mathrm{D} = c\mathrm{T}.$$

This equation does not fix the position of the ship but only shows that it lies on a circle whose centre is at the station and whose radius is D. If, however, the same signals are received at the second station, distant d from the ship, and the time interval was t, the position of the ship must be on a circle of radius d. By drawing two circles of radii D and d respectively and with centres corresponding to the positions of the stations the position of the ship could be determined uniquely.

The distance of the ship from each of the land stations having been ascertained by them the information is sent by wireless to the ship, the total time elapsing since the ship informed the land stations that a knowledge of its whereabouts was required being about ten minutes.

The Multiple-Charge Method for determining the Velocity of Sound in Sea-Water.—This method was developed about 1923 by A. B. Wood and others in conjunction with the Admiralty. Let us suppose that B, Fig. 31·12 (*a*), is a buoy fixed in position approximately in line with two hydrophones (instruments for detecting sound waves in water) H_1 and H_2, at a known distance apart, D, but at a considerable distance from H_1. Now suppose that a destroyer steering a straight course at uniform speed along a line XY passing through B and at right angles to $\mathrm{H}_1\mathrm{H}_2$, drops at approximately equal intervals a number of depth charges, about half of these lying on either side of B.

H_1 and H_2 are connected to a land-station so that a film record may be made of the time of arrival of sound signals at the above two points. Now the time of passage of the explosion-wave across the base line, $\mathrm{H}_1\mathrm{H}_2$, will increase from charge No. 1 to charge No. 6 and then diminish. If the time intervals are plotted with respect to the time of dropping the charges from the destroyer, the points will

lie on a smooth curve with a well-defined maximum—cf. Fig. 31·12 (*b*)—between the 6th and 7th charges in the particular instance here recorded. If t_m is the maximum time interval for the passage of the wave between H_2 and H_1, then

$$c = \frac{D}{t_m}.$$

In deriving this formula the only assumptions made are that the

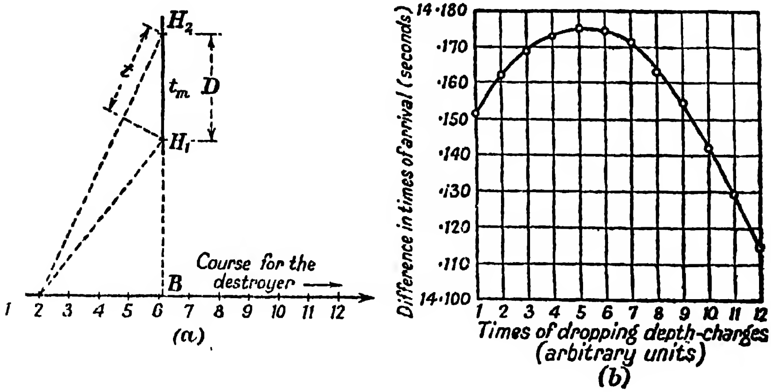


FIG. 31·12.

speed of the destroyer is approximately constant and that its course is approximately a straight line.

In some of these experiments H_1H_2 = 70,245 ft.; t_m = 14·176 sec. Hence c = (4955·4 ± 1) ft. sec.$^{-1}$ at 17° C. Tidal errors were eliminated by carrying out the experiments at a time of neap-tide slack water. The film record enabled the time differences to be measured accurately to within 0·002 second. The twelve depth charges were fired at intervals of 1¼ min. (± 2 sec.), and the course of the destroyer was kept straight by steering directly towards a fixed landmark.

Sound-ranging on Land.—During the First World War the need for a method of locating the position of an enemy gun became very urgent. The following method was therefore developed. Suppose S, Fig. 31·13, was a source of sound—the gun—while A, B, and C, were three observers who recorded the time when a sound from the gun reached them. If the flash of the gun had been seen, the method of deducing the gun's position was exactly the same as that used in finding the position of the ship as described in the previous section—moreover, only two stations would have been essential. The presence of the third station was due to the fact that very often the flash was not seen. At each station there was concealed a piece of very thin platinum wire supported on a mica frame as in Fig. 31·13 (*b*) and mounted in front of a resonator. [This consists of a large vessel containing air which responds loudly to notes of certain pitch which are always present in the array of sounds caused by the explosion of a gun, but are absent from the sounds due to speech, traffic, etc.—cf. p. 652.] The wire was placed in one of the arms of a Wheatstone bridge and the bridge balanced in the usual way. When the resonator was in action the violent excursions of the air past the microphone,

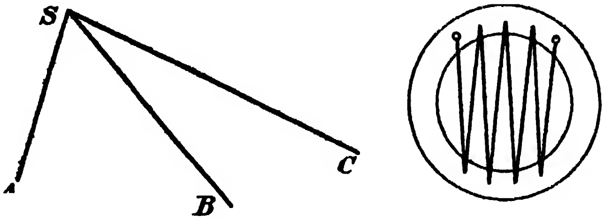
i.e. the wire, mounted as described above, cooled it and the bridge became unbalanced since the resistance of platinum decreases when the temperature is lowered. This want of balance in the bridge was recorded electrically at the base which was connected to each station. If c is the velocity of sound in air under the prevailing atmospheric conditions, and t_a, t_b, and t_c the times when the signal was received at each station

$$\frac{SA}{c} - \frac{SB}{c} = t_a - t_b$$

and

$$\frac{SB}{c} - \frac{SC}{c} = t_b - t_c.$$

These equations indicate that S lies at the intersection of two hyperbolæ, having their foci at A and B, and at B and C, and in which



(a) (b)

FIG. 31·13.—Sound Ranging.

the differences between the focal distances of any point on the curves are $c(t_a - t_b)$, and $c(t_b - t_c)$ respectively. These hyperbolæ were constructed by a specially designed curve tracer and the position of the gun was found. As the method is employed at present the reception of the sound is indicated by a slight protuberance on an otherwise straight line on a photographic film. This line is produced by light reflected from the mirror of a galvanometer used to determine when the bridge is balanced. Each protuberance shows that the balance has been disturbed. The film is fed continuously into highly concentrated developing and fixing solutions so that the negative may be examined within a few seconds of the time when the signal was received. From the speed at which the film is fed into the camera the time intervals can be derived.

EXAMPLE XXXI

1. Assuming the adiabatic relation between pressure and volume for a perfect gas, deduce an expression for the adiabatic elasticity. Why is this value used in calculating the velocity of sound in a gas ?

2. Describe and explain an accurate method for measuring the velocity of sound in air.

3.—What is the experimental evidence for the view that sound consists of waves propagated through a material medium ? How does the velocity of the waves through air depend on their length, and on the pressure and temperature of the air ?

CHAPTER XXXII

REFLEXION, REFRACTION, AND INTERFERENCE OF SOUND WAVES

The Characteristics of Sounds.—The sounds to which our ears respond may be divided into two classes:—(i) Sounds of short duration which change their character continually if they persist for some time; they are termed ***noises***; (ii) Sounds which are characterized by their smoothness and regular flow, as distinct from the irregularity and impulsive nature of noises, and termed musical sounds. Musical sounds or notes may differ from one another in three important particulars: they may differ (*a*) in ***intensity***, i.e. in ***loudness***; (*b*) in ***pitch***; (*c*) in ***quality*** [or ***timbre***].

Intensity-Loudness.—This depends upon the amount of energy carried by the incident waves and is analogous to the brightness of a source in optics. The ***intensity*** of a sound is measured by the amount of energy passing per second through an area 1 cm.^{2} drawn perpendicular to the direction of propagation at the point concerned.

If a particle of a mass m is moving with velocity u, its kinetic energy is $\frac{1}{2}mu^2$. If the particle is executing simple harmonic motion represented by $y = \hat{a}\cos(\omega t - \phi)$ [cf. p. 580], its velocity at any instant is

$$u = \dot{y} = -\hat{a}\omega\sin(\omega t - \phi).$$

$$\therefore \text{Kinetic Energy} = \text{E} = \tfrac{1}{2}m\,\hat{a}^2\omega^2\sin^2(\omega t - \phi)$$
$$= \tfrac{1}{4}m\hat{a}^2\omega^2[1 - \cos 2(\omega t - \phi)].$$

The second term in the bracket ranges in value from $+1$ to -1, its average value over a complete period being zero, since it is just as often positive as negative. The average value of E is therefore $\frac{1}{4}m\,\hat{a}^2\omega^2 = \frac{1}{2}\cdot\frac{1}{2}m\hat{U}^2$, where $\hat{U} = \hat{a}\omega$, i.e. it is half the maximum kinetic energy.

Now the energy passing through an area 1 cm.^{2} at any particular point per second, the area being at right angles to the line of propagation, is equal to that of all the particles in a column of area 1 cm.^{2} and length c where c is the velocity of the sound. The mass of the column is ρc, where ρ is the density of the medium. The energy proportional to the intensity of the sound is therefore $\text{W} = \frac{1}{4}\cdot\rho c\cdot\hat{a}^2\omega^2$, an expression which shows that the intensity is

proportional to the square of the amplitude. Now loudness is a physiological effect depending on the ear of the listener, but the more intense the sound, the more loud does it appear. By the method adopted on p. 351 it can be shown that the intensity I is inversely proportional to the square of the distance, r, from the source to the point considered, so that

$$I \propto \hat{a}^2 \propto r^{-2}.$$

Whence $\hat{a} \propto r^{-1}$, i.e. the amplitude, is inversely proportional to the distance from the source.

Pitch.—Pitch in acoustics corresponds to colour in optics: in fact, pitch may be referred to as musical colour. It is determined by the frequency of the vibrations and increases with increase in frequency. Now whereas musical notes are characterized by a definite and constant frequency all noises are an array of notes of varying pitch, i.e. of varying frequency.

Savart's Wheel.—When a piece of thin metal sheet is held against the teeth of a rotating wheel the former executes an impulsive vibratory motion in consequence of the regular impacts it receives from the wheel. A note is emitted when the number of impacts per second is sufficiently great, and the frequency of the note increases with the speed of rotation.

The 'Cardboard' Siren.—This consists of a circular cardboard or metal disc capable of revolving about an axis passing through its centre and normal to its own plane, and having a number of equidistant holes drilled near to its periphery. A jet of air impinges upon the disc which is arranged so that the holes are close to the jet. When a hole comes in front of the jet a puff of air passes through. As the disc rotates a series of puffs is produced and if these succeed one another sufficiently rapidly a note is produced. The pitch of the note rises as the speed of the disc increases.

Quality.—Musical notes may have the same pitch and intensity, and yet differ from one another: they are said to be of different ***quality*** or ***timbre.*** To explain this we have to remember that the note emitted, when any one note of a piano is struck, is seldom pure; e.g. a trained ear is able to detect the presence of notes which are higher than that of the ***fundamental*** or main note. These higher notes, if they have frequencies which are low integral multiples of that of the fundamental are known as ***overtones***; other notes which may be present are termed ***upper partials.*** These higher notes cause a change in the wave-form of a note and give to it a certain distinctiveness or timbre. It is because the voices of our acquaintances differ in timbre that we are able to distinguish one from another.

Sensitive Jets and Flames.—To examine the progress of a liquid emerging from a jet into water the apparatus shown in Fig. 32·1 may be used. A 'constant head' reservoir of the type already described [cf. p. 313] is connected to a tube AB about 2 cm.

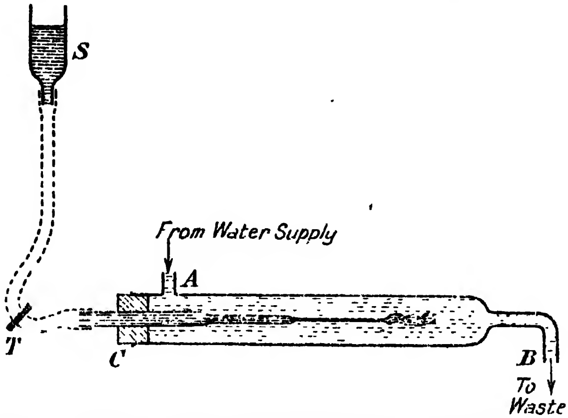


FIG. 32·1.—Osborne Reynold's Experiment on Turbulent Motion.

in diameter, the liquid entering at A and escaping at B. At the end C is a rubber bung carrying an inlet tube drawn out to a capillary about 10 cm. long and 0·5 mm. in diameter. This is joined to a small reservoir S filled with ink. The flow of ink is controlled by a spring clip T. The position of the reservoir is adjusted by means of the flexible tubing connecting it to A until a long column of ink is seen escaping from the jet. When these conditions have been attained the liquid is moving with a ***stream-line motion.*** [This experiment succeeds best when the flow of ink is not too rapid.] By increasing the rate of flow of the water the stream-line motion is destroyed and the jet of ink moves irregularly. The motion has become ***turbulent.***

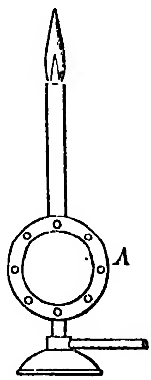


FIG. 32·2.—Rayleigh's Sensitive Flame.

Similar happenings take place when a jet of coal gas, for example, escapes into the atmosphere; they may be followed by igniting the gas.

Experiment.—A piece of glass tubing 0·5 cm. in diameter, drawn out to form a jet 0·5 mm. in diameter, is connected to a coal-gas supply and placed about 3 cm. below a piece of fine copper gauze arranged horizontally. The gas above the gauze is ignited, but the flame does not extend below the copper. The position of the jet is adjusted until the flame is on the point of flickering. When a high

note is sounded in its neighbourhood the flame ducks violently; similar happenings take place when a bunch of keys is rattled near to the flame.

Another form of apparatus for producing a sensitive flame is due to Lord Rayleigh. A brass cylinder A, Fig. 32·2, about 4 cm. long and 5 cm. in diameter is closed at one end, the other being covered with a piece of thin tissue paper or mica. Gas entering from below passes through the cylinder and is ignited at the top of the exit tube [15 cm. long]. The gas flow is adjusted until the flame is apparently detached from the apparatus, when it will be found sensitive to various sounds, especially if they are of an explosive nature like the letters *p*, *b*. This flame is exceptionally responsive to high notes if the orifice is covered with a cap pierced with a small hole.

Reflexion of Sound.—Two cardboard tubes about 10 cm. in diameter and 100 cm. long and inclined to each other are placed in front of a cardboard or other screen, S, Fig. 32·3. A watch, or a Galton's whistle [cf. p. 636] is placed at A and a sensitive flame at B. The screen R serves to prevent sound waves reaching the sensitive flame directly: to effect this more completely it is better to surround A entirely by a screen. The screen S is rotated slowly until the flame indicates that sound-waves are incident upon it. If the angles between the axes of the tubes and the normal to S at N are measured they will be found to be equal. It may also be shown that sound-waves are reflected from concave mirrors in the same manner as are light and heat waves by placing a source of sound at the focus of one mirror and a sensitive flame at that of the other. If the two mirrors face each other the sensitive flame responds when in this position. The experimental arrangement is similar to that described on p. 328.

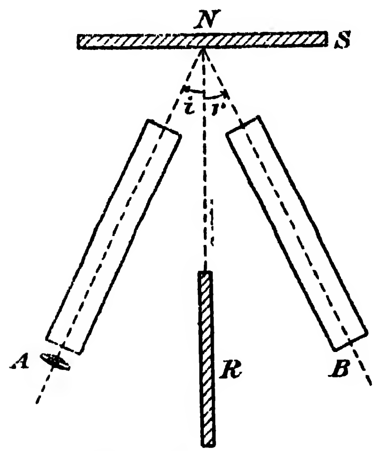


FIG. 32·3.—Reflexion of Sound Waves.

The reflexion of sound-waves at the face of a mountain cliff is responsible for the formation of echoes: whispering galleries owe their peculiar properties to this same phenomenon. The reflexion of sound from the walls of some buildings is a cause of much annoyance—this was especially so in the House of Commons. It is now known that these echoes may be minimized by avoiding sharp corners and covering the walls of the room with 'acoustic plaster' which absorbs much of the incident sound energy.

Refraction of Sound.—The following experiment is due to TYNDALL :—A large soap bubble is blown with carbon dioxide and placed between a source of sound and a sensitive flame. The flame responds most readily for one particular position for a fixed position of the source with respect to the soap bubble. This is because the heavier gas in the bubble causes the sound to be refracted and the flame is affected most strongly when it and the source occupy positions known as conjugate foci.

That sound-waves do suffer refraction is easily demonstrated by the following experiment :—A sensitive flame is placed at some distance from a source of sound—the flame responds. A coil of resistance wire is placed between them and heated electrically. The response of the flame is less vigorous.

Refraction by a Wind.—It is a well-known fact that a sound travelling with the wind is better heard than when it travels in the reverse direction.

An explanation of this was first given in 1857 by STOKES. Let us suppose that a source of sound-waves is at S, Fig. 32·4 (*a*), and

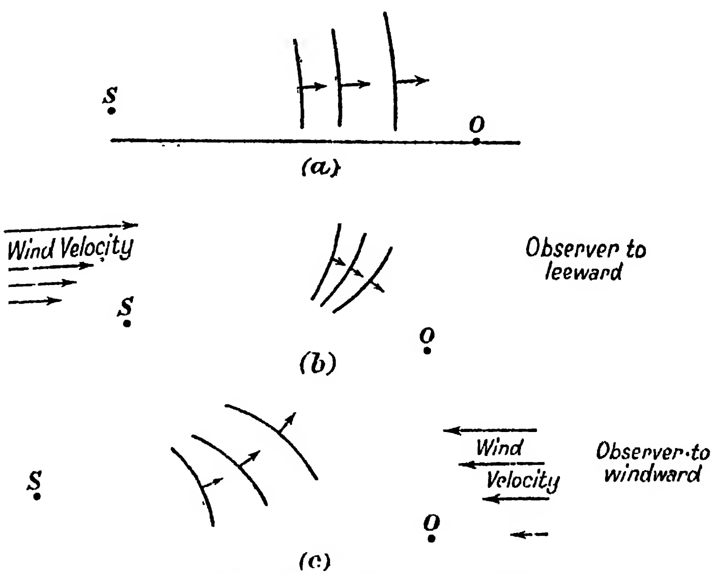


FIG. 32·4.—Refraction by the Wind.

that O is an observer on the ground. The wave-surfaces will be spherical, becoming less curved as they advance. When there is a wind the layers of air will themselves move so that a disturbance moves with a velocity which is the algebraic sum of the velocity of sound in still air and that of the wind. Thus the velocity of

a disturbance will be greater or less than that in air at rest according as the sound travels with the wind or against it. Let us first consider the effect when the sound travels with the wind, i.e. the observer is on the leeward side of the source. Fig. 32·4 (*b*) shows that the wave-fronts will move as if they were proceeding from a source above the ground, i.e. the portions of the waves reaching the observer are those which initially tended to travel upwards: they have had the advantage of being out of the way of obstructions for the greater part of their course. On the other hand if the observer is on the windward side of the source, as in Fig. 32·4 (*c*), the upper portions of the wave-fronts will be retarded with respect to the lower ones, so that the path along which the energy flows will gradually be bent upwards, and at a moderate distance passes over the observer.

Similar conditions to the first occur even if there is no wind when the temperature increases upwards. The velocity of sound-waves then increases with height [cf. p. 593], so that on both sides of the source the wave-fronts are bent downward and thus the sound will be better heard by an observer. If the temperature decreases with height the sound is not heard so well.

Reflexion of Sound-waves by a Wall.—Let PQ, Fig. 32·5, represent a rigid wall upon which a train of sound-waves is falling at normal incidence. Since the wall is rigid, none of the energy incident upon it can be transmitted forward so that the layer of air in contact with the wall must remain permanently at rest, for if it moved away there would be a vacuum on one side and a pressure almost atmospheric on the other. Hence, when a compression reaches the wall, the only way whereby this layer can free itself from its strained condition is by pushing back its neighbours. A compression is therefore reflected from the wall as a compression. Similarly, a rarefaction is reflected as a rarefaction. In each case the displacements of the particles due to the incident wave are opposite to those due to the reflected wave. The waves are said to have been reflected with ***change of phase*** since the motions of the particles are reversed by reflexion. To discover the state of the air through which these two trains of waves are passing we have to construct the velocity or displacement curves and find their combined effect. The reflected waves can be represented by a wave-train moving from left to right. Since the velocities of the air particles adjacent to the wall are always zero, the velocities due to each train must be equal and opposite at this point, i.e. at the wall the two wave-trains differ in phase by π at every instant. The reflexion is said to have occurred with change of phase. In the diagrams the thin and dotted lines represent the velocity curves of the incident and reflected waves respectively. The thick line

represents their resultant. At time $t = 0$ let us assume that the velocity of the incident wave due at the wall is a maximum ; the velocity at this point due to the reflected wave will also be a maximum equal but opposite in direction to the above. At this instant all the particles have a resultant velocity equal to zero [cf. Fig. 32·5 (*a*)]. At time $\frac{T}{8}$ seconds later the curves shown in (*b*) represent the state of affairs. The thin curve is really the thin

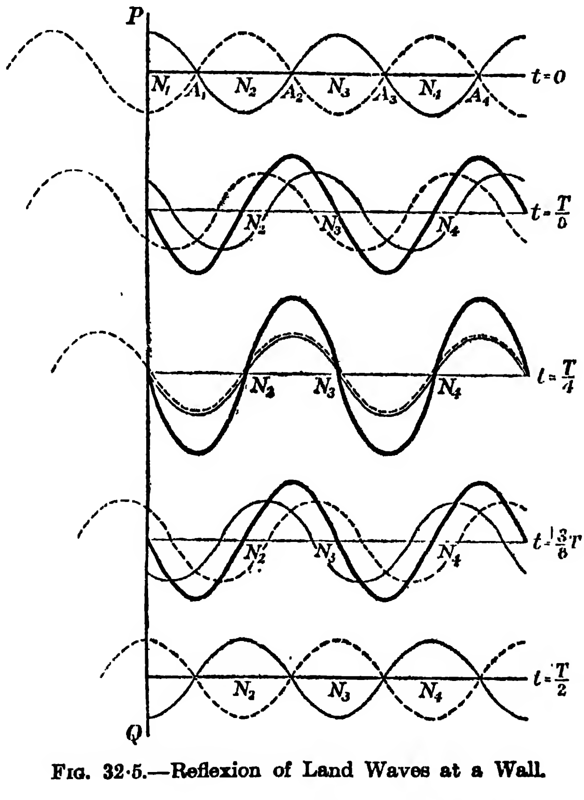


FIG. 32·5.—Reflexion of Land Waves at a Wall.

curve in (*a*) displaced one-eighth of a wave-length to the left, while the dotted one is obtained from (*a*) by advancing the thin curve one-eighth of a wave-length to the right. The resultant is shown by the heavy line. The curves in (*c*), (*d*), and (*e*) have been drawn in the same manner.

In a single wave-train the amplitude is the same for each vibratory particle although the maximum displacement of each occurs at different times. In the present example the points N_1, N_2, N_3, N_4, etc., are permanently at rest. They are termed ***nodes*** and are

separated from each other by an amount $\frac{\lambda}{2}$, where λ is the wave-length of the motion. Moreover, the particles in between the nodes have different amplitudes although the amplitude is a maximum at the same instant. The points where the amplitude is a maximum e.g. at A_1, A_2, A_3, etc., are called ***antinodes.***

Vibrations similar to this are termed ***stationary vibrations*** or ***standing waves.*** In our treatment above it has been assumed that the reflecting wall was perfectly rigid. In actual examples some of the incident energy will pass into the medium of which the wall is a boundary so that the amplitude of the reflected waves will generally be less than that of the incident ones. In virtue of this the velocity and displacement at the nodes is never exactly zero.

Experiment.—In high-frequency sound-waves, such as are produced when a Galton's whistle is blown, the positions of the nodes and anti-nodes may be located with the aid of a sensitive flame. When the flame is at a node it does not flicker since the molecules are at rest, but when at an antinode violent flickerings of the flame manifest themselves since at these points the disturbances are most pronounced. Since the distance apart of successive nodes or antinodes is $\frac{\lambda}{2}$, the frequency of the note may be determined if the velocity of sound in air is known. In this manner the frequency of a note which is so high that it is beyond the upper limit of audibility may be found, and since the frequency of the note emitted by a Galton's whistle is continuously variable the upper limit of audibility may be fixed.

Reflexion without Change of Phase.—If sound-waves travelling in the more dense of two media impinge upon a boundary between them reflexion takes place under conditions very different from those just discussed. Let us suppose that a wave of amplitude a is advancing from the right towards the wall PQ, Fig. 32·5. As the wave passes along, each layer acquires energy which is expended in imparting motion to the next layer. When the layer adjacent to PQ is set in motion it retains some of its energy when it has moved through a distance a since the particles to the left of PQ are more easily set in motion. In consequence of this, the layer continues to move towards the left until it has advanced a total distance $b[b > a]$. This causes the air behind to become rarefied: thus a reflected wave of rarefaction of amplitude b is set up by the compression wave incident on the boundary PQ. This is termed a ***reflexion without change of phase.*** At the interface between the media the displacements are large so that antinodes occur here. The first node is therefore at a distance $\frac{\lambda}{4}$ from the interface.

Interference of Sound-waves.—We have already discussed the conditions under which light-waves may interfere. Since sound is a wave-motion [differing from light in that it consists of longitudinal waves transmitted through a material medium and that its wave-length is considerably greater] it follows that sound-waves should be capable of interfering. Now with light-waves interference can only occur if the interfering trains have their origins in the same source; interference may be shown with sound-waves emitted from different sources.

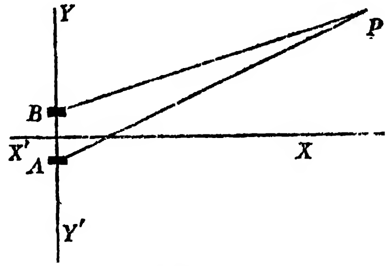


FIG. 32·6.—Interference from Two Sources.

Let A and B, Fig. 32·6, be two sources emitting sound-waves of the same amplitude and frequency. At the point P [not necessarily in the plane of the diagram] the two wave-trains will reinforce each other if they are in phase, so that an observer at this point will hear a loud sound. On the other hand, if the waves differ in phase by $\frac{\lambda}{2}$ or, in general, by $(2n + 1)\frac{\lambda}{2}$ where n is any integer, the medium will remain undisturbed since the displacements due to each set of waves are equal and opposite. These effects will persist as long as the sources continue to vibrate since we have supposed that their periods are equal.

Conditions for Interference.—The following conditions must be fulfilled if two wave-trains are to interfere with each other:—

(*a*) The frequencies of the waves must be the same, otherwise any difference in phase at a particular point would not be maintained, and, if mutual destruction occurred at one instant, reinforcement would take place soon afterwards.

(*b*) The amplitudes [i.e. intensities] of the two vibrations must be equal, otherwise complete interference is impossible. If the amplitudes are not equal the positions in which the phase difference is $(2n + 1)\frac{\lambda}{2}$ will not be positions where the resultant displacement is zero.

(*c*) The displacements should be collinear, for, otherwise, the motion of the particles at points where the phase difference is $(2n + 1)\frac{\lambda}{2}$ would not be zero and the particle at each such point would execute a type of Lissajous' figure.

Quincke's Tube.—The wave-length of a high-frequency note may be determined with the aid of QUINCKE's tube shown in Fig. 32·7. This consists essentially of two tubes A and B about 3 cm. in diameter and bent as indicated. The effective length of the right-hand tube may be altered by sliding the tube A. Let us suppose that a Galton's whistle is blown near to C. The sound-waves entering the tube may travel to D via the path CAD or CBD. If these are equal the two sets of waves will be in phase when they reach D so that if a sensitive flame is placed at this point it will be violently disturbed. On the other hand, if the tube A is moved, a position will be reached when the two trains differ in phase by $\frac{\lambda}{2}$ when they arrive at D. When this occurs the flame will not flicker,

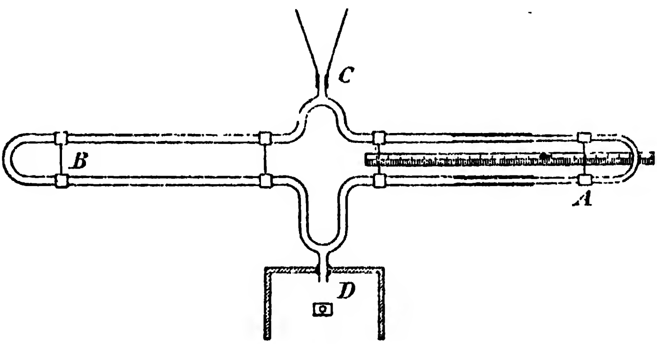


FIG. 32·7.—Quincke's Tube.

and the tube A will have been withdrawn a distance $\frac{\lambda}{4}$. When the tube A is moved through a distance $\frac{\lambda}{2}$ the path difference will be λ, so that the waves at D will be in phase and the flame flicker, but this will cease if the displacement is increased to $3\frac{\lambda}{4}$. Proceeding in this way several positions of the tube A may be found such that destructive interference occurs at D. The distance between any two consecutive positions is $\frac{\lambda}{2}$, so that the wave-length and hence the frequency may be determined.

Beats.—Suppose that one of two tuning-forks whose frequencies are identical is loaded with a small quantity of wax so that its frequency is diminished. If the two are sounded together it will be noticed that the resultant intensity waxes and wanes. These alternations of strong and weak sounds are termed ***beats***. If the difference in frequency is n, then, at any point, n times every second

the phases of the waves will be the same and f times per second they will differ by π, i.e. by an amount equivalent to half a wave-length.

Exercise.—Construct, after the manner indicated in Chap. XXXI, two sine curves in which the frequencies are 8 and 9. Then construct the resultant curve formed by adding the two motions together. In this curve it will be noticed that the amplitude fluctuates in a perfectly regular manner. The maximum disturbance at any given point occurs when the two separate disturbances are in phase, while, when the one wave-train has made one half-vibration more than the other, the two will be in opposite phase and the resulting disturbance a minimum—if the amplitudes are equal, the resultant amplitude is zero.

By counting the number of beats per second when two forks are sounding we determine at once the difference in frequency between two notes. To determine which is the fork of lower frequency one of them is loaded. If the number of beats per second is increased then the fork which is loaded had the lower frequency originally, for the load has merely served to increase the difference in frequency between the forks. Care must be taken to repeat the experiment with the other fork loaded when the number of beats per second should be reduced. It is necessary to do this, for it is possible that the higher-frequency fork may be so heavily loaded that the number of beats per second is increased instead of being diminished as it would be if the load were not too great.

Beats.—*Analytical Treatment.* Let the two wave trains to be compounded have amplitudes $\hat{a}_1$ and $\hat{a}_2$, and differ slightly in frequency (or wave-length), i.e. they are given by the equations

$$\begin{aligned}\xi_1 &= \hat{a}_1 \cos(\omega t - \phi_1),\\ \xi_2 &= \hat{a}_2 \cos[(\omega + \Delta\omega)t - \phi_2],\\ &= \hat{a}_2 \cos[\omega t - (\phi_2 - \Delta\omega . t)].\end{aligned}$$

The resultant is therefore given by

$$\begin{aligned}\xi = \xi_1 + \xi_2 &= \hat{a}_1 \cos\omega t \cos\phi_1 + \hat{a}_1 \sin\omega t \sin\phi_1\\ &\quad + \hat{a}_2 \cos\omega t \cos(\phi_2 - \Delta\omega . t) + \hat{a}_2 \sin\omega t \sin(\phi_2 - \Delta\omega . t)\\ &= \hat{a}\cos(\omega t - \phi) \qquad \text{(say)},\end{aligned}$$

where

$$\begin{aligned}\hat{a}\cos\phi &= \hat{a}_1\cos\phi_1 + \hat{a}_2\cos(\phi_2 - \Delta\omega . t)\\ \hat{a}\sin\phi &= \hat{a}_1\sin\phi_1 + \hat{a}_2\sin(\phi_2 - \Delta\omega . t),\end{aligned}$$

i.e.

$$\hat{a} = [\hat{a}_1^2 + \hat{a}_2^2 + 2\hat{a}_1\hat{a}_2\cos\{\phi_1 - (\phi_2 - \Delta\omega . t)\}]^{\frac{1}{2}},$$

and

$$\tan\phi = \frac{\hat{a}_1\sin\phi_1 + \hat{a}_2\sin(\phi_2 - \Delta\omega . t)}{\hat{a}_1\cos\phi_1 + \hat{a}_2\cos(\phi_2 - \Delta\omega . t)}.$$

Hence the amplitude fluctuates between $(\hat{a}_1 + \hat{a}_2)$ when the cosine term is $+1$, and $(\hat{a}_1 - \hat{a}_2)$ when it is -1. The period, τ, of these fluctuations is such that if

$$\begin{aligned}\phi_1 - (\phi_2 - \Delta\omega . t_1) &= \alpha, \qquad \text{say},\\ \phi_1 - \{\phi_2 - \Delta\omega(t_1 + \tau)\} &= \alpha + 2\pi\end{aligned}$$

for then the displacements are identical. Thus

$$\tau = \frac{2\pi}{\Delta\omega} = \frac{2\pi}{2\pi f_1 - 2\pi f_2} = \frac{1}{f_1 - f_2},$$

where f_1 and f_2 are the frequencies.

Thus the resultant vibration is one whose period is $\frac{2\pi}{\omega}$, i.e. it is equal to that of the disturbance ξ_1, but whose amplitude fluctuates: it is a maximum $(f_1 - f_2)$ times per second and this is equal to the number of beats heard per second.

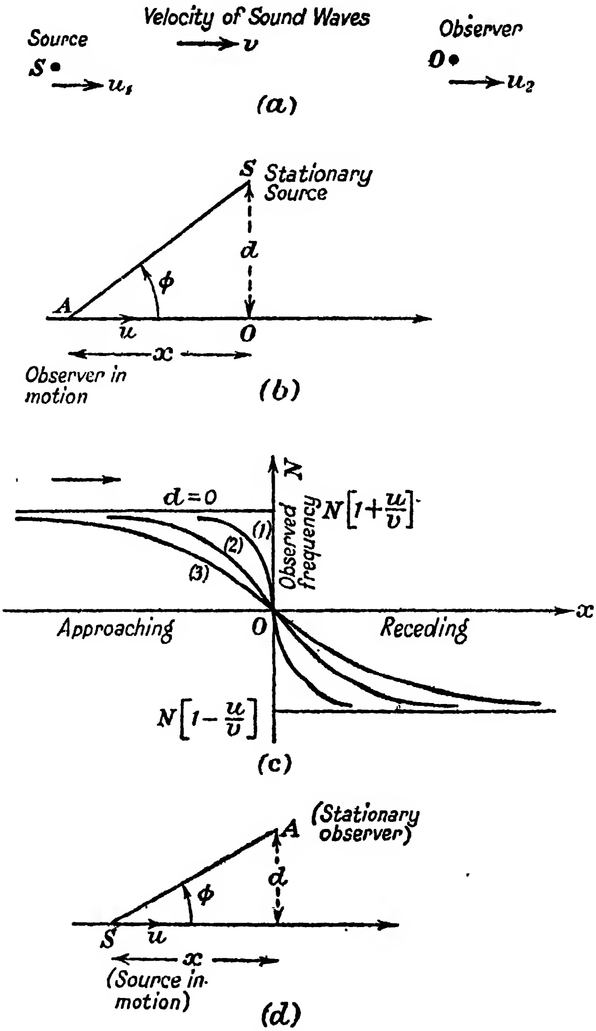


FIG. 32·8.—The Doppler Effect.

The Doppler Effect (or the Variation of Pitch with Motion). —Let S, Fig. 32·8 (*a*), be a source of sound-waves, the frequency being f. Suppose this source moves with velocity u_1 towards an

observer O who is moving away from the source with a velocity u_2, each of these velocities being relative to a fixed point. Then the disturbed wave-length, λ', is given by

$$\lambda' = \frac{\text{Velocity of separation of wave-fronts and source}}{\text{Frequency of emission from source}},$$

$$= \frac{c - u_1}{f}. \qquad [c = \text{velocity of sound.}]$$

The frequency of the waves as received by the observer is given by

$$\frac{\text{Velocity of approach of wave-fronts and observer}}{\text{Disturbed wave-length}} = \frac{c - u_2}{\lambda'}$$

$$= f.\frac{c - u_2}{c - u_1}.$$

[If there is a wind blowing with velocity v_w along the line SO the velocity c must be increased by v_w — it must be diminished if the wind is in the opposite direction.]

Now let us consider the variation in pitch caused by motion between a source and observer when the motion is not along the line joining them. Thus in Fig. 32·8 (*b*) let A be an observer moving with velocity u along a straight line. Let S be the stationary source at distance d from the above line. Let $\widehat{SAO} = \phi$. Then the observer is approaching S with a velocity $u \cos \phi$, so that the frequency of arrival of the waves at P, which is the apparent frequency of the source, is given by

$$f' = \frac{\text{Velocity of approach of wave-front and observer}}{\text{Disturbed wave-length}}.$$

$$= \frac{c + u \cos \phi}{\lambda} \quad \text{[Since in this instance the disturbed wave-length is equal to the actual wave-length } \lambda.]$$

$$= f.\frac{c + u \cos \phi}{v}$$

$$= f\left[1 + \frac{u}{c}\cos \phi\right] = f\left[1 + \frac{u.x}{c(d^2 + x^2)^{\frac{1}{2}}}\right], \text{ where AO} = x.$$

In Fig. 32·8 (*c*) the curves indicate how the value of N′ varies with x for different values of d. When $d = 0$ N′ has a constant value on either side of the origin, viz. $f\left[1 + \frac{u}{c}\right]$ on the 'approaching' side and $f\left[1 - \frac{u}{c}\right]$ on the 'receding' side. Thus, when the observer passes through O, when $d = 0$, there is a sudden change in pitch $= \frac{2fu}{c}$.

Again, let the observer be at rest at A, Fig. 32·8 (d), while the source, S, is in motion with velocity u along a straight line whose distance from A is d. Now in the case just considered, i.e. when the observer was in motion and the source was at rest, the apparent change in frequency was due to the change in the rate at which the sound-waves passed the observer—there was an apparent change in the velocity of propagation, the wave-length remaining constant. In the instance now under discussion there is an apparent change of wave-length due to the motion of the source towards the observer, the velocity of sound remaining unchanged. The disturbed wavelength is given by

$$\lambda' = \frac{\text{Velocity of separation of wave-fronts and source}}{\text{Actual frequency}}$$

$$= \frac{c - u\cos\phi}{f}.$$

$\therefore$ Frequency of arrival at A

$$= \frac{\text{Velocity of approach of waves and observer}}{\text{Disturbed wave-length}},$$

$$= \frac{c}{c - u\cos\phi}.f = \frac{f}{\left[1 - \dfrac{ux}{c(d^2 + x^2)^{\frac{1}{2}}}\right]}.$$

When $d = 0$, the change in pitch as the source moves past the observer is from $\dfrac{f}{\left(1 - \dfrac{u}{c}\right)}$ to $\dfrac{f}{\left(1 + \dfrac{u}{c}\right)}$. For other values of d and x, curves similar to those in Fig. 32·8 (c) could be drawn.

In all cases it will be noted that for small values of d, i.e. when the source and observer are actually close together at some instant, the apparent frequency is at first constant and then suddenly drops to a constant lower value. As d increases the change is more gradual so that for large values of d the change is less marked, unless u is also large. These last remarks are illustrated by the following examples: (1) When an engine passes an observer close to the railway line there is a sudden drop in the pitch of the note from the whistle if this is sounding; (ii) the passage of an aeroplane overhead.

The apparent change in frequency due to relative motion between a source of waves and an observer was first discussed and elucidated by Doppler. He had noticed the apparent change in wave-length of the lines in stellar spectra. According to his principle, when a star is approaching the earth, the light waves will be shorter, i.e. the lines in the spectrum of the light from the star will be shifted towards the region of shorter wave-lengths—viz. towards

the violet end of the spectrum. Similarly, a shift towards the red end of the spectrum will denote a motion away from the earth. Other examples of the use to which the Doppler principle has been put are as follows: (i) by examining the spectrum of the light from the periphery of the sun, its speed of rotation has been calculated; (ii) Saturn's rings have been shown to be rotating more rapidly at the inner edge than at the outer.

The Reflexion of Plane Waves (at Normal Incidence) from a Moving Reflecting Plane (or Mirror).—Let a train of plane waves be incident normally on a plane 'mirror' M, Fig. 32·9 (*a*). The waves travel from left to right. Let λ_0 be their wave-length and c their velocity. Let A_0, A_1, A_2 . . . represent successive crests in the advancing wave-train. Now suppose that M moves towards the oncoming waves with velocity u. Let us consider the state of affairs from the instant when A_0 just reaches M. Suppose

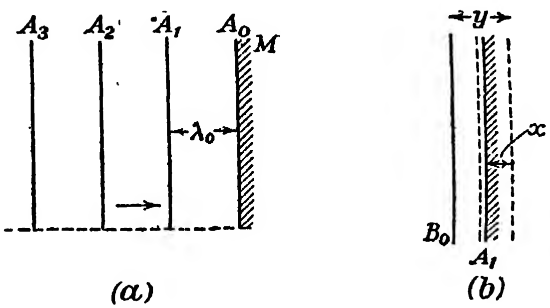


FIG. 32·9.

that the mirror meets A_1 when it has moved forward a distance x from the chosen zero position—cf. Fig. 32·9 (*b*). Then A_1 and M have moved relatively a distance λ_0 with a velocity $(c + u)$. This will occur in time $\dfrac{\lambda_0}{(c+u)}$. During this time the reflected crest corresponding to A_0 will have advanced to B_0, i.e. through a distance y equal to $c.\dfrac{\lambda_0}{(c+u)}$, while the mirror will have advanced to A_1 through a distance $\dfrac{\lambda_0}{(c+u)}.u = x$. But B_0A_1 is the wave-length of the reflected train, viz.

$$(y - x) = \frac{c-u}{c+u}.\lambda_0 = \lambda_1 \text{ (say)}.$$

If $\dfrac{u}{c} \to 0$ (as is always the case when light waves are considered)

$$\lambda_1 = \left(1 - \frac{2u}{c}\right)\lambda_0.$$

If the mirror moves in the same direction as the oncoming waves

$$\lambda_{\mathrm{ref.}} = \frac{c - (-u)}{c + (-u)}.\lambda_0 = \frac{c + u}{c - u}.\lambda_0,$$

which is $\left(1 + \frac{2u}{c}\right)\lambda_0$, when $\frac{u}{c} \to 0$.

EXAMPLE XXXII

1.—**Explain the alteration of the pitch of a note with the motion of the source. An engine travelling at 60 ml. hr.$^{-1}$ passes an observer at rest. If 580 cycle. sec.$^{-1}$ is the frequency of the note heard when the engine sounds its whistle while moving towards the observer, what is the frequency of the note which may be heard when the engine is receding ? [Velocity of sound in air = 1100 ft. sec.$^{-1}$.]**

2.—A line of wave-length $4{\cdot}8 \times 10^{-5}$ cm. in the spectrum of the light from a star is found to be displaced from its normal position nearer to the red end of the spectrum by an amount corresponding to a change in wave-length of $1{\cdot}2 \times 10^{-8}$ cm. What velocity of the star in the line of sight would account for this shift ?

3.—Two trains are approaching one another with speeds of 60 and 45 km.hr.$^{-1}$. A whistle of frequency 512 cycle.sec.$^{-1}$ is sounded on the first train. Calculate the frequency of the note heard by an observer on the second train (*a*) before and (*b*) after the two trains pass one another.

CHAPTER XXXIII

THE VIBRATIONS OF STRINGS, RODS, AND COLUMNS OF GAS

The Velocity of Transverse Waves along a Stretched String.—For our present purpose a string may be defined as a perfectly flexible uniform filament of cord or wire. Since all actual strings possess rigidity it is necessary to use thin strings when designing experiments to check our theoretical deductions since, in thin strings, the effect of rigidity is a minimum. Let us assume that a perfectly flexible string, having a mass μ per unit length, is stretched by a force F (absolute units). To deduce the velocity of transverse waves along such a string we shall use the method originated by TAIT. He imagined that the string was passed

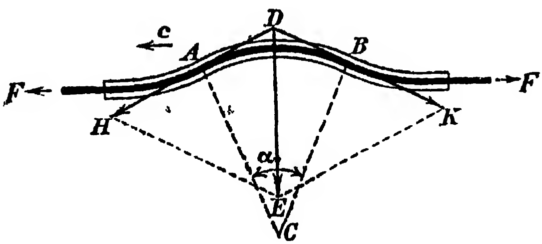


FIG. 33·1.—Velocity of Transverse Waves along a String.

through a smooth tube of the shape shown in Fig. 33·1, with a velocity c. The tension in the string gives rise to a pressure tending to straighten the tube and string, whereas the tube tends to increase the curvature of the string. When these two effects are equal and opposite the form of the curved portion of the string remains stationary in space, each portion of the string assuming this shape in turn. Relative to the string, the curved portion moves with a velocity c. When these conditions have been attained, consider a portion AB, of length l. If the tensions at the ends of AB are represented by DH and DK respectively their resultant, P, which is the force exerted on the tube round AB,

is represented completely by DE. If a is the angle between normals AC and BC, then $DE = 2.DH \sin\frac{a}{2}$; hence $P = 2F \sin\frac{a}{2}$, or Fa, when α is small. Now the centrifugal force due to a mass μl moving with a speed c in an arc whose radius is r, is $\frac{\mu l c^2}{r}$. When this is equal to the force due to the tension in the cord, the tube may be removed, and the velocity of the transverse motion is given by

$$F.a = \frac{\mu l c^2}{r}.$$

But $\alpha = \frac{l}{r}$, so that $c = \sqrt{\frac{F}{\mu}}$.

The Transverse Vibrations of Strings.—The strings dealt with in practice are always of finite length and attached at each extremity to a rigid support, so that when a disturbance reaches the extremity of the string a reflected wave will be set up. Since this wave is reflected at a rigid wall there will be a change in phase $\frac{\pi}{2}$, so that a node is always found at the end of the string. The simplest possible type of vibrating string is one in which there are only two nodes, i.e. the length of the string is one-half the wave-length of the disturbance travelling along it. If l is the length of the string, then $\lambda = 2l$. Since f the frequency of the vibration is expressed by $f\lambda = c$, we have

$$f = \frac{c}{2l} = \frac{1}{2l}\sqrt{\frac{F}{\mu}}.$$

The expression just obtained may be verified experimentally by means of a ***sonometer*** or ***monochord***—Fig. 33·2. This instrument is said to have been in use at the time of Pythagoras but the elementary laws of vibrating strings were not made known until 1636. In that year MERSENNE expressed them separately. The formula given above was established theoretically by BROOK TAYLOR in 1715.

Mersenne's Laws.—These laws were formulated as the result of experiments on the vibrations of strings long before the mathematical theory had been developed. They are :—

(i) ***For a given string and given stretching force the frequency varies inversely as the length.***

(ii) ***For a uniform string of given length and material, the frequency varies as the square root of the stretching force.***

(iii) ***The frequency of vibration of strings of the same length and subjected to the same stretching force varies***

inversely as the square root of the mass per unit length of the string.

Lamb in his book on sound writes: 'The principles that the frequency diminishes with increase of length and with increase of line-density have a familiar illustration in the pianoforte, where longer and intrinsically heavier strings are used for the graver notes. If the relation of pitch were adjusted by length alone the strings corresponding to the lower notes would be at least 100 times as long as those belonging to the highest. In order to secure a sufficiently low pitch within practical limits of length, and with a sufficient degree of tension, the string is loaded with a coil of wire wrapped closely round it. This has the effect of increasing the inertia without seriously impairing the flexibility, which is an essential point. The influence of tension, again, is illustrated in the process of tuning, which consists in tightening up the wires when these have been stretched, or the pegs have yielded, so that the instrument has fallen in pitch, or become "flat."'

The Sonometer.—The sonometer consists of a wooden board or box upon which two wires are stretched. One of the wires, cf. Fig. 33·2, is stretched by means of a mass supported by it over a pulley. The other is placed under tension by wrapping it round an iron peg which may be rotated by means of a wrench key, as are the wires in a piano. The vibrations are confined to definite portions of the wires by means of two fixed bridges A

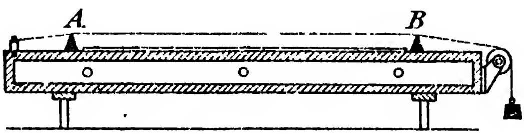


FIG. 33·2.—A Sonometer.

and B. Other small movable bridges are supplied so that any length of wire can be selected for use. The wooden box is a desirable feature since it vibrates in tune with the wire. The mass of air affected is greatly increased in this way so that the loudness is augmented. [Does this violate the principle of the conservation of energy? No, for the vibrations die away much more rapidly than when the wire alone vibrates.] The wooden body of a violin, and the sounding board of a piano, behave in an analogous manner.

***Experiment.**—To show that* $f \propto \frac{1}{l}$. Attach a constant load to the wire passing over the pulley, and adjust one of the movable bridges until the wire vibrates in unison with a tuning fork of known frequency. Students having difficulty in judging the equality of two notes may obtain the final adjustment with the aid of a wooden disc about 4 in.

in diameter attached to one end of a short wooden rod whose axis is normal to the plane of the disc. The free end of this rod is placed in contact with the sonometer board while the disc is pressed against the ear. The tuning fork is struck and held against the board. If the adjustment is approximately correct, beats will be heard. The bridge is moved until the beats are very slow, when the length of the vibrating wire is recorded. The observations are repeated with other forks. Since theory shows that fl is constant, the verification of this fact may be shown by plotting $\log f$ against $\log l$, when a straight line having a slope -1 should be obtained.

Experiment.—To show that $f \propto F^{\frac{1}{2}}$. A sonometer, with two stretched wires, is required. One of these wires is kept under a constant load, and its length is adjusted so that the frequency of the wire when plucked shall be equal to that of a given tuning fork. [Strictly speaking, it is not necessary to know the actual value of the frequency of this fork.] This wire now furnishes us with a scale of frequencies, for if its length is altered, its frequency will be changed in such a way that the product fl is constant.

Now let a fixed length of the second wire be selected and let it be under a tension F. The frequency of the note it emits when plucked is found by adjusting the length of the first wire until there is resonance between the two wires. A series of corresponding values of F and f having been obtained, $\log f$ is plotted against $\log F$ and if a straight line whose slope is 0·5 is obtained, the fact that $f \propto F^{\frac{1}{2}}$ will have been verified.

Experiment.—To show that $f \propto m^{-\frac{1}{2}}$. Again a sonometer with two wires is required. One of them is kept under a constant load and standardized as above so that it provides us with a scale of frequencies. A selected length of the second wire is then put under a constant load and the frequency determined with the aid of the standardized wire. Its mass per unit length, m, is then found. An equal length of another wire, of different mass per unit length, is then placed under the same load and its frequency determined. A series of such readings should be obtained. If $fm^{\frac{1}{2}}$ = constant, then

$$\log f + \tfrac{1}{2} \log m = \text{const.}$$

By plotting $\log f$ against $\log m$ and obtaining a straight line whose slope is $-\frac{1}{2}$, the fact that $f \propto m^{-\frac{1}{2}}$ will have been verified.

On Tuning Two Notes to Unison.—The student who has no musical 'ear' will have difficulty in deciding when the frequencies of two notes are equal. If an attempt is made to tune a fork and string to unison, and the string is in a horizontal position, the following method may be adopted to indicate when the tuning is correct. A small paper rider is placed at the middle of the string and the sounding fork allowed to rest on the board of the sonometer. When the tuning is approximately correct the rider will flutter, and will be thrown off when the fork and string are in unison. This occurs because the fork and string are in such a condition that when one is sounding the other resounds. The experimental procedure, therefore, is to vary the length of the

string so that the fluttering increases and the rider is eventually thrown off.

The above method may also be used when two strings attached to the same sonometer board are to be tuned to unison.

Another method of deciding when two notes are in accord is as follows. Its applicability is of a more general nature than the above. When two notes have approximately the same frequency beats will be heard; by adjusting the frequency of one of the notes the beats are made so slow that they cannot be distinguished. The tuning is then exact, i.e. the frequencies of the two notes are identical.

To Determine the Absolute Frequency of a Tuning Fork.—The sonometer wire, stretched by a known load, is tuned until it is in unison with the fork. The equality may be tested by listening for beats in the manner already described. The frequency is then calculated from the formula

$$f = \frac{1}{2l}\sqrt{\frac{F}{\mu}}.$$

Unless a thin string is used this frequency will not be the true frequency of the fork, for the rigidity of the wire will produce an extra force tending to restore the wire more quickly to its zero position when vibrating, i.e. the frequency will be increased.

In all work with sonometers it must be remembered that if M is the mass of the load carried by the wire the stretching force is not Mg, where g is the intensity of gravity, on account of friction at the pulley. It is therefore necessary to write $F = Mg + C$, where C is a constant, so that

$$f = \frac{1}{2l}\sqrt{\frac{Mg + C}{\mu}},$$

which may be written

$$\frac{4\mu f^2}{g}l^2 = M + \frac{C}{g}.$$

If, therefore, we write $l^2 = x$, $y = M$, we obtain a straight line whose slope is $\frac{4\mu f^2}{g}$ if C is strictly a constant. It is most likely to be constant in comparison with the other terms if it is small: hence the pulley wheel should not be too small in diameter.

The above equation also enables us to determine the density of the material of a wire if a standard fork is available. The wire is adjusted until in tune with the fork, when the mass per unit length of the wire may be determined. If the density of the material of the wire is ρ, and r is its radius, $\mu = \pi r^2 \rho$.

The Frequency of an Alternating Current.—A small current from a source of alternating current is sent along a sonometer wire, AB, Fig. 33·3, the central portion of which lies between the opposite poles S_1, N_2 of two cobalt steel magnets $N_1 S_1$, $N_2 S_2$, i.e. this portion of the wire is in a strong magnetic field, if the above poles are near together. The tension in the wire is adjusted until resonance occurs, i.e. the wire vibrates vigorously since its

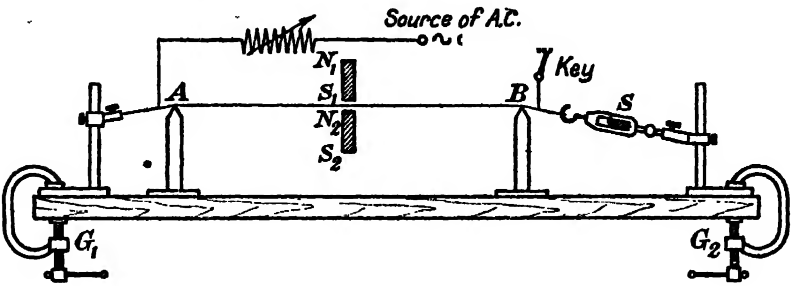


Fig. 33·3.—To measure the Frequency of an Alternating Current Supply.

own natural period is the same as that of the alternating current. The frequency, f, is then calculated from the equation

$$f = \frac{1}{2l}\sqrt{\frac{F}{\mu}},$$

where l is the length of wire between the bridges of the sonometer, μ the mass per unit length of the wire, and F the tension in the wire. S is a spring balance which measures F. G-clamps, G_1 and G_2, prevent the clamps supporting the wire from falling when the latter is under tension.

Again, on account of friction, a series of observations with different loads should be made, when, as on p. 621, we have

$$f = \frac{1}{2l}\sqrt{\frac{Mg + C}{\mu}},$$

so that f may be determined graphically.

Harmonics and Overtones.—The vibrations of a wire so far considered have been such that there have been only two nodes present. The wire has then given its lowest or *fundamental* note. It can be made, however, to vibrate so that intermediate nodes exist. For example, if four nodes are to appear it is only necessary to touch the wire lightly with the aid of a feather at a point distant one-third of the length of the wire from one end, and to bow the wire with a violin bow at an antinode. The first four modes of vibration of a stretched wire are indicated in

Fig. 33·4. If the fundamental is Doh, the notes emitted when three, four, and five nodes are present are doh, soh, and doh′ respectively.

The presence of these nodes may be made apparent by placing small paper riders on the wire which is then bowed, while a feather touches the wire at a point $\frac{1}{n}$th of its length from one end, where n is a small integer. The riders will be thrown off at the antinodes or loops, but will remain on the wire at the nodes.

When a stretched wire vibrates a well-trained musical ear detects notes of a frequency different from that of the fundamental note. These higher frequencies are called *overtones*. When the wire is stretched between rigid supports it is found, both theoretically and in practice, that all the overtones are integral multiples of the fundamental frequency. Such overtones are known as ***harmonics***, the fundamental frequency being called the first harmonic, the first overtone (with a frequency twice that of the fundamental) is the second harmonic, the second overtone is the third harmonic, and so on.

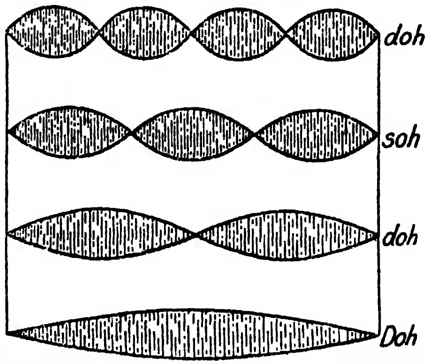


FIG. 33·4.—The First Four Modes of Vibration of a Stretched String.

Harmonic overtones are not found in many vibrating systems, but those systems where they do exist are used in nearly all musical instruments, for the sounds caused by such are pleasant to a musical ear.

The presence of overtones due to a piano wire may easily be shown as follows:—A piano key somewhere near the centre of the board is pressed down and held in that position. When the note has died away the key an octave below is struck vigorously and then released. The first wire will be heard vibrating. This is because it has picked up notes having the same frequency as those it emits when vibrating. These were present in the vibrations of the second wire and constitute its first overtone or second harmonic.

The Experiments of Melde.—A very beautiful method of demonstrating the vibrations of stretched strings is due to MELDE. An electrically maintained tuning fork [p. 627] is clamped to a table as in Fig. 33·5. One end of a string about two metres long is attached to one prong of the fork while the other end is joined

to a pan after passing over a pulley—arranged in this way, the direction of motion of the point of attachment is parallel to the length of the string. The pan is loaded, the fork excited, and the load, or length of string, adjusted until the vibrating string shows one loop. If the load is reduced to one quarter the above value, two loops will be obtained; when it is reduced to one-sixteenth, four loops will be produced. G-clamps serve to hold the apparatus in position. If for any pattern thus obtained the plane of

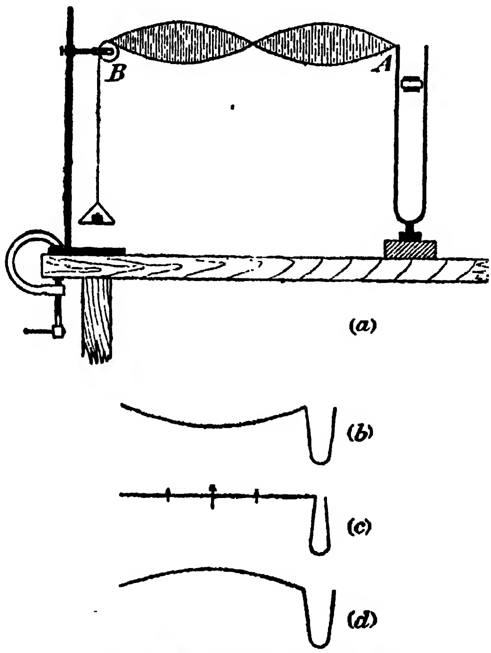


FIG. 33·5.—Melde's Experiment.

the fork is rotated through 90°, everything else being kept the same, the wire will be found to be vibrating with twice the number of loops.

The explanation of these phenomena may be obtained by considering Fig. 33·5 (*b*), (*c*), and (*d*). In the first instance when the point of attachment moves parallel to the undisturbed position of the string, the effect of the motion causes the tension in the string to vary periodically so that, at first sight, there is apparently no reason why the string should depart from its equilibrium position of straightness. Actually, the equilibrium of such a system can be shown to be unstable, and the string settles down in a state of vigorous vibration with a period twice that of the fork, i.e. its

frequency is one-half that of the fork. A general explanation is as follows. When the prong A has made its maximum excursion towards B, let us assume that the amount of sag in the string is also a maximum. As the prong returns the sag decreases, becoming zero when the prong has made its maximum excursion to the other side. When the prong returns the wire does not sag but is carried upward in virtue of the inertia it possesses. When the prong has reached the position it formerly had in (*b*), the wire is at rest at its maximum displacement above the horizontal—cf. (*d*). This shows that to every complete vibration of the fork in this position the string makes one-half of a complete vibration. The frequency of the string is therefore one-half that of the fork.

When the motion of the prong is at right angles to the string, i.e. when the fork has been rotated through 90°, the string will move to the right when the prong moves to the right; it will be at rest when the prong is at its zero position; it will move to the left when the prong moves to the left; i.e. the vibrations of the string will synchronize with those of the fork and the two frequencies will be equal.

The Transverse Vibrations of Rods.—Our considerations of the vibrations of strings have been made on the assumption that the strings are perfectly flexible, i.e. the strings are restored to their zero positions after being displaced solely in virtue of the tension in them. The opposite extreme is that of a vibrating rigid rod. Here there is no tension along the rod and the restitution is brought about by the stiffness of the material of the rod. The vibrations of a rod fixed at one end executing its fundamental and first two nodes of vibration in addition to the fundamental are indicated in Fig. 33·6, but a full treatment of the subject shows that they are not exact harmonics of the fundamental.

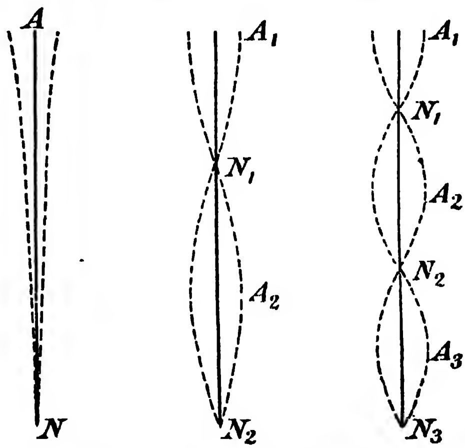


FIG. 33·6.—Transverse Vibration of a Rod fixed at One End.

Tuning Forks.—When a solid rod is supported at two points, N_1 and N_2, Fig. 33·7 (*a*), and caused to execute transverse vibrations, these two points become nodes—there may be other nodes

intermediate between these and on each side of them, but the important feature about the motion is that the two ends are always moving in the same direction at the same time. When the bar is sounding its fundamental there are only two nodes.

If a rod is gradually bent at its centre, the two nodes, when the bar is sounding its fundamental, approach the centre of the bar as the bending increases—cf. Fig. 33·7 (*b*). When the two portions of the rod are parallel the nodes are very close together

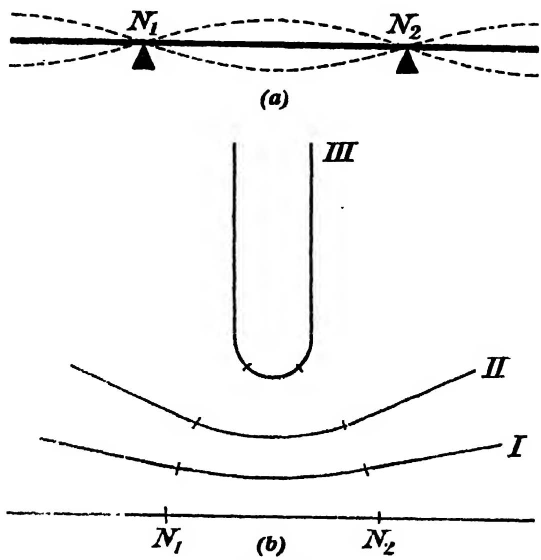


FIG. 33·7.

and the motion is that of a tuning fork. This method of examining the nature of the vibrations due to a fork gives us the reason why the prongs of a fork always approach or recede from each other.

Tuning forks play an important part in the study of sound because a properly designed and constructed fork furnishes us with a ready means of obtaining a note which is practically free from overtones providing it is not bowed too vigorously. It is always difficult to excite these overtones, and even when they are produced they are very feeble and die away much more rapidly than does the fundamental.

Electrically Maintained Forks.—Sometimes, however, it is necessary to have a fork which shall emit a note continuously for some time. For low notes, when the prongs are long and heavy, the vibrations may last for a minute or more, but the higher notes

from forks having short prongs die away much more rapidly. In either instance they may be maintained electrically as follows:—The tuning fork is rigidly mounted in a brass collar A, Fig. 33·8, while an electromagnet, B, is placed symmetrically between its free ends. One prong carries a small platinum style which rests in contact with a platinum disc C when the fork is silent. C is attached to a screw so that its position with respect to the style may be varied by rotating the head of the screw. T_1 and T_2 are

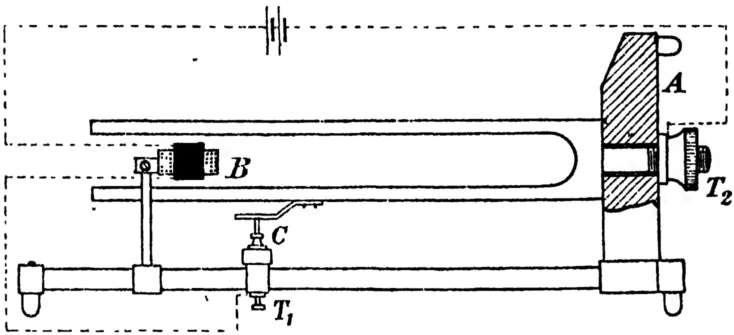


FIG. 33·8.—Electrically Maintained Tuning Fork.

two terminals. The electrical circuit is completed as indicated by dots, the current passing (with the usual convention with respect to its direction) from the battery through the electromagnet to T_1; from thence to C and through the fork to the pillar A and the terminal T_2. When this is occasioned the magnet is excited and attracts the prongs of the fork, thereby breaking the circuit. The prongs then move back, contact is made at C, and the whole process continued.

The Longitudinal Vibrations of Rods.—Solids, in addition to executing transverse vibrations when suitably stimulated, may, like gases, execute longitudinal vibrations. The frequency of such vibrations is independent of the tension along the rod, for when a particle is temporarily displaced from its position of rest the forces tending to restore it arise in virtue of the elasticity of the material of the rod. We have already stated that the speed of longitudinal waves is given by $c = \sqrt{\frac{E}{\rho}}$, where E is Young's modulus and ρ the density of the material through which the waves are propagated.

When a rod is clamped at its centre there must be a node at this point, and when the fundamental is being sounded the free ends must be antinodes or loops. The wave-length of the sound in the rod will be twice the length of the rod since the distance from

node to loop is one-quarter of a wave-length and this is half the length of the rod in the present instance.

Resonance.—When a body whose *natural frequency* is f_1 is subjected to a periodic force having a frequency f_2, the resulting motion depends upon how nearly the *impressed frequency* f_2 equals f_1. Let us assume that a pendulum, initially at rest, and whose natural frequency is one per second, is subjected to a succession of small blows at intervals of 1·01 seconds. This constitutes an intermittent impressed force having a frequency $\frac{100}{101}$. After the first blow, the pendulum begins to move with its own frequency, but when it receives the second blow it will have made more than one complete vibration and be moving in the direction along which the impressed force acts. Consequently its momentum will be increased so that it moves beyond its initial maximum displacement. This process will continue for some time, the amplitude being increased after each blow. At the twenty-sixth blow the pendulum will have made $25\frac{1}{4}$ complete oscillations, i.e. it will receive the blow at an extreme position. At the twenty-seventh the pendulum will be moving in a direction opposite to that in which the blow is struck so that its amplitude begins to decrease. Gradually the pendulum will be brought to rest, after which, the whole cycle of events will be repeated.

When the difference between the natural frequency of the fork and that of the blow gets less, the pendulum will execute more complete oscillations before the blow begins to reduce the amplitude of its swing. Meanwhile, if the magnitude of the blow remains constant, the amplitude will have continued to increase after each blow. In the limit, when the two frequencies are equal, the magnitude of the oscillation would become infinite.

When the impressed force is periodic instead of being intermittent, the body subjected to its influence may be set in a periodic motion. When the period of the impressed force is the same as the natural period of the body very energetic oscillations of the latter will be produced. This is an example of *resonance*. If the natural period of the body does not agree with that of the applied force the vibrations set up will, in general, be of small amplitude and the period will equal that of the impressed force. When a body is performing vibrations not agreeing with its own natural period the vibrations are said to be *forced*. On removing the impressed force the body will continue to vibrate in virtue of the inertia it possesses, but the period will be equal to the natural period of the body.

A study of resonance phenomena is of great importance to the engineer, for, if the period of even a small impressed force agrees

with that of the body to which it is applied, the amplitude may attain such values that a fracture ensues. It is for this reason that a regiment of soldiers always breaks step when crossing a bridge. Similarly the effect of resonance may be very pronounced in ships fitted with reciprocating engines. If the period of the reciprocating masses is identical with the natural period of the hull the amplitudes of the motion of the latter may become dangerous. It is therefore essential to see that the two periods do not coincide.

Some Examples of Resonance and Forced Vibrations.—(*a*) Let two forks of equal pitch and mounted on their resonance boxes be so placed that the open ends of the latter face each other. If, after one has been bowed and allowed to sound for a short time, it is stopped, the second fork will be heard although initially it was silent.

(*b*) Support an indiarubber tube AB, about one metre long, at its ends as indicated, Fig. 33·9, and suspend three simple pendulums CD, EF, and HG from points C, E, and H respectively. Adjust CD and HG so that their lengths are equal. Pull CD *forwards,* so that when released it vibrates in a plane at right angles to that of the diagram. The tube AB acts as an intermediary for the transmission to the other pendulums of the energy due to the vibrating

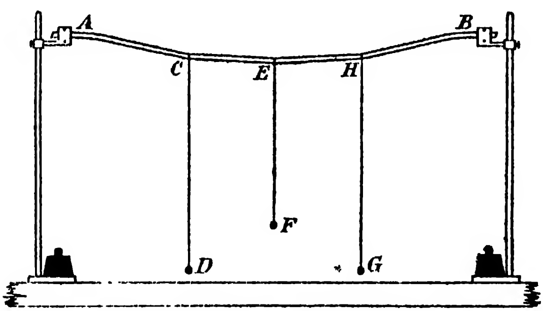


FIG. 33·9.—Simple Pendulums in and out of Resonance.

pendulum CD. Since the periods of the two pendulums CD and HG agree the amplitude of the latter will increase until it has absorbed all the energy initially possessed by CD [except for small inherent losses]. CD will then be at rest and HG vibrating with an amplitude practically equal to that of CD originally. CD will then begin to receive energy and its amplitude increase until HG has been reduced to rest; the process continues until the energy has all been dissipated as heat in overcoming frictional and other resistances. The pendulum EF, having a period different from those of CD and HG, only executes forced oscillations of hardly perceptible amplitude. This is all the more remarkable when EF is placed, as in the diagram, between the other pendulums.

(c) A remarkable instance of forced vibrations occurs when two clocks which keep approximately the same time when placed on different stands maintain the same time when on the same stand. The pendulum of the clock which gains normally exerts a periodic force on the second so that the two periods tend to become equal; the second clock exerts a similar effect on the first so that eventually the two periods are equal and the clocks synchronize.

Vibrating Columns of Gas.—Columns of gas enclosed in tubes of uniform bore may be caused to vibrate longitudinally in a manner exactly analogous to rods executing longitudinal vibrations. Two types of gas column present themselves: (a) when the containing tube is closed at one end—the so-called ***closed tube,*** and (b) when the containing tube is open at both ends. This latter is termed an ***open tube.***

Let AB, Fig. 33·10 (a), be a tube closed at one end. Let the length

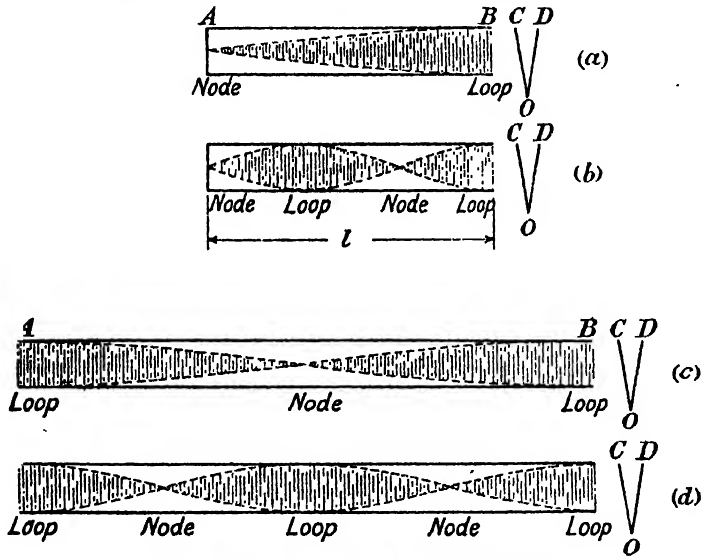


Fig. 33·10.—Resonance of Air Columns.

[It must be pointed out that the vibrations are ***longitudinal*** and not ***transverse*** as shown for convenience in the diagram.]

of this tube be equal to one-quarter the wave-length of the note emitted by a given tuning fork. If this vibrating fork is held at the mouth of the tube, then when the prong of the fork is about to leave the position OD and travel towards OC [greatly exaggerated in the diagram] a compression just begins to pass down the tube. This compression travels to the end A where it is reflected as a compression [cf. Chap. XXXII]. When this compression reaches the

open end of the tube the prong of the fork is just about to return from OC to OD, so that the layers of air immediately outside the tube are more readily moved: the compression passes outwards. A rarefaction then begins to move down the tube and this in turn is reflected from A. When this arrives at B the external air moves toward the rarefied layers and a compression is sent down the tube. Since the length of the tube is such that the time for a wave to travel from B to A and return again is $\frac{T}{2}$, where T is the period of the fork, the compression sent down the tube after the first rarefaction has left will begin its journey at the instant when the fork itself is sending a compression down. The compression due to the reflexion at the open end and that due to the sounding fork will be identical in phase so that the column of gas in the tube will be caused to undergo violent ***stationary vibrations*** of the same frequency as the fork. There will be a node at A and an antinode at B. If $AB = l$, the waves travel a distance $2l$ in time $\frac{T}{2}$, so that its wave-length, λ_1, the distance travelled in time T, is $4l$. The frequency, f_1, of the fork is $\frac{c}{4l}$. When a column of air vibrates in sympathy with a fork it is said to be in ***resonance*** with the fork, The same column of air may also be in resonance with a fork of higher frequency f_3 if the length l is such that $l = \frac{3}{4}\lambda_3$, i.e. $f_3 = \frac{3c}{4l}$, cf. Fig. 33·10 (*b*). Similarly if $l = \frac{5}{4}\lambda_5$, the tube will respond to a note of frequency $f_5 = \frac{5c}{4l}$. When the tube is open at both ends, as in Fig. 33·10 (*c*), the simplest possible longitudinal vibration which can arise will have a node at the centre of the tube and two antinodes, one at each end. In this instance a compression is reflected from A as a rarefaction which is then returned from B as a compression. If the period of the fork is such that a wave travels from B to A and back again in time T, then the compression from B due to reflexion, and the direct compression due to the fork will begin to travel down the tube together; since they are in phase the vibrations of the tube will become vigorous. The same tube can also respond to another fork if its length is such that stationary waves having three antinodes and two nodes as in Fig. 33·10 (*d*) are produced. Using the same notation as before

$$l = \frac{\lambda_1}{2}, \text{ i.e. } f_1 = \frac{c}{2l}. \quad \text{Similarly } f_2 = \frac{2c}{2l}, f_3 = \frac{3c}{2l}, \text{ etc.}$$

These equations indicate an important difference between the fundamental and overtones produced with columns of gas in open

and closed tubes. In the first instance the only overtones are odd harmonics of the fundamental, while in the second all the harmonics may be present as overtones [cf. p. 623]. Hence although an open and a closed pipe may be made to emit the same fundamental note the quality will be very different in the two instances.

The Measurement of λ by Resonance Tubes.—The apparatus, Fig. 33·11 (*a*), consists of a tube AC about 5 cm. in diameter and a metre long. It is connected at its lower end by means of rubber tubing to a reservoir R containing water (or better a light oil having a negligible vapour pressure). The reservoir is raised until the water stands near to the top of the tube. The clip E is adjusted so that when the reservoir is lowered the water flows slowly from the tube. While this is happening a sounding fork is held over the tube. When the water-level in the tube is at some particular and well-defined position the air in the tube responds to the vibrations of the fork. Let B be this position. Now the length AB is not exactly $\frac{\lambda}{4}$ since the simple theory developed above is only approximate. We assumed that the open end is an antinode. Lord Rayleigh showed that the antinode is situated at a short distance outside the tube. The magnitude of this ***end correction*** is $0{\cdot}58r$, where r is the radius of the tube. Actually there is no need to assume the value of this correction for the real wave-length and the end correction may be determined as follows :—

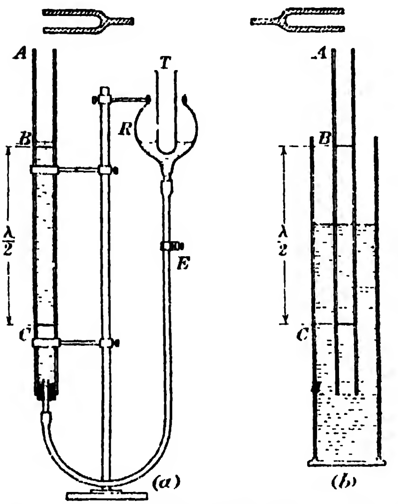


Fig. 33·11.—Closed Resonance Tubes.

Water is allowed to escape from the tube until it responds again to the fork—say at C. Then if l_1 and l_2 are the lengths AB and AC respectively, we have

$$(l_1 + \theta) = \frac{\lambda}{4} \text{ and } (l_2 + \theta) = \frac{3\lambda}{4}, \text{ etc.}$$

where θ is the correction in cm. Hence $(l_2 - l_1) = \frac{\lambda}{2}$. These equations enable both λ and θ to be obtained. The end correction, in terms of r, is then deduced.

If desired, one may dispense with the clip E, and when a position of resonance has been located approximately, the water-level in AC may be caused to change slowly by raising or lowering a boiling tube, T, placed in the reservoir R as shown.

As an aid to locating the positions at which the tube responds to a given fork, i.e. the tube 'speaks,' the fork should be moved slowly in a horizontal plane across the mouth of the tube. When the length of the air column in the tube is appropriate, the response of the tube is very noticeable.

Another form of apparatus often used in this connexion is shown in Fig. 33·11 (*b*).

Open Resonance Tubes and the Determination of λ.—AB, Fig. 33·12 (*a*), is a glass tube open at both ends. Its length must be less than $\frac{\lambda}{4}$, where λ is the wave-length to be determined. C is a cardboard tube sliding over AB. The given fork, while sounding, is held near one end of the tube—or better, moved slowly across, as above—and the amount by which C projects adjusted until the tube 'speaks.' The length of the tube is $\left(\frac{\lambda}{2} - 2\theta\right) = l_1$, say, where θ is the correction for each end of the tube. D is a second cardboard tube—see Fig. 33·12 (*b*)—and it is adjusted until the tube again speaks. The total length of the tube is then $(\lambda - 2\theta) = l_2$, say. The difference $l_2 - l_1$ is $\frac{\lambda}{2}$.

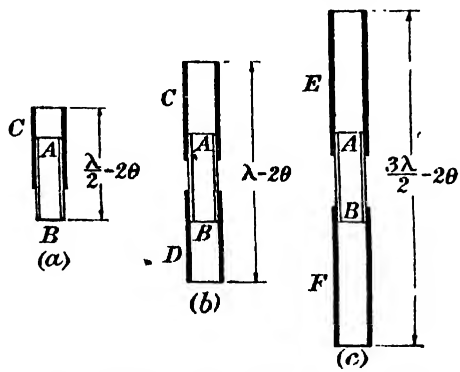


FIG. 33·12.—Open Resonance Tubes.

Fig. 33·12 (*c*) shows the next position of the tubes E and F (they may have to be longer than C and D) when the tube speaks. The total length is $\left(\frac{3\lambda}{2} - 2\theta\right) = l_3$, say, and $l_3 - l_1 = \lambda$.

In experiments, such as the above, with tubes, it must be remembered that we assume that the velocity of sound in free air, where it is measured directly, is the same as in air confined in a tube. Strictly speaking, this can hardly be expected to be so, for true adiabatic conditions cannot possibly exist in the tube at points close to the wall. Moreover, it should be verified experimentally that the wave-length of a note of given frequency in free air is

equal to the distance between one node and the next-but-one node in air when stationary waves have been produced.

Experiment.—Examine the effect on θ of covering the end of a closed resonance tube by sheets of copper in each of which a hole, of different diameter from the rest, has been made.

Experimental Study of the Effect of Temperature on the Velocity of Sound in Air.—The effect of change in temperature on the velocity of sound in air is an important phenomenon already treated theoretically on p. 593. The following apparatus, due to Bateman, allows us to investigate the problem experimentally. It consists of a metal tube (105 cm. × 3·5 cm.) AB, Fig. 33·13 (*a*), provided with a steam jacket, the steam entering and escaping from suitably placed side tubes. The length of the air column is varied by adjusting the position of a piston P. The tube is supported in a slightly inclined position on metal legs provided

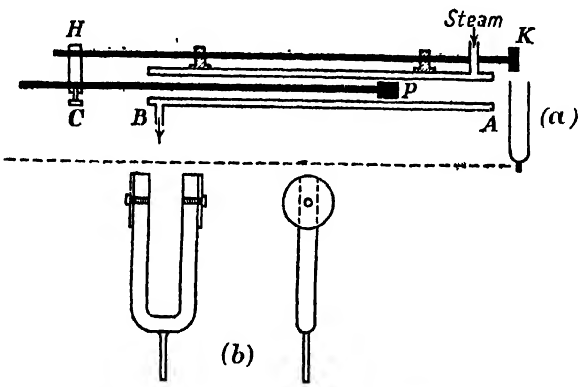


Fig. 33·13.—Effect of Temperature on the Velocity of Sound in Air.

with rubber feet, the open end being lower than the rest of the tube so that convection currents may not be unduly harmful. The tuning fork used is provided with two metal discs—Fig. 33·13 (*b*) —each about the size of a shilling and screwed to the prong of the fork. This alters the frequency of the fork, but this quantity is not essential unless absolute values of the velocity of sound in air at different temperatures are required.

The tube is first used at room temperature and the first and second positions of resonance obtained. The wave-length of the disturbance in the air is deduced in the usual way—let it be λ_1. The piston is first adjusted approximately from the end B of the tube when the rod attached to it is clamped by C to another rod HK which slides in supports attached to the steam chamber and

which is operated by the head K (made of heat-insulating material). The final adjustment is made from the open end of the tube, the sounding fork being held in position.

The experiment is then repeated with steam passing through the jacket. Let λ_2 be the wave-length under these conditions. If c_1 and c_2 are the velocities of sound in air at temperatures t_1 and t_2,

$$\frac{c_2}{c_1} = \frac{\lambda_2}{\lambda_1}.$$

If, therefore, it is found that $\frac{\lambda_2}{\lambda_1} = \sqrt{\frac{273 + t_2}{273 + t_1}}$, the fact that the velocity of sound in air is directly proportional to the square root of the absolute temperature will have been verified.

Organ Pipes.—These are wooden or metal tubes having a square or circular cross-section. A 'stopped diapason,' an organ pipe of wood and of rectangular section, is indicated in Fig. 33·14. The wind at a constant pressure of several inches of water passes into the mouthpiece, M, and escapes from the linear slit O. It then impinges upon the edge E formed by bevelling the wall of the pipe. An adjustable piston S closes the pipe whose 'speaking length' is from S to a point somewhere in the neighbourhood of O. The air blast, on striking E, gives rise to 'edge-tones.' If the length of the tube is such that the tube responds readily to one of these tones, the tube 'speaks.' The movement of the air is a maximum at the mouth so that this becomes an antinode. The other end becomes an antinode or node according as the pipe is 'open' or 'stopped.' The simple theory we developed in connection with vibrating columns of gas does not apply in this instance owing to uncertainties regarding the end correction at the lip. The tuning must therefore be done experimentally. This is accomplished in closed pipes by varying the position of a movable piston which serves to close the tube. With open pipes the tuning is done by raising or lowering a flap placed at the open end of the pipe so that the end correction is altered: this causes a change in the pitch of the pipe.

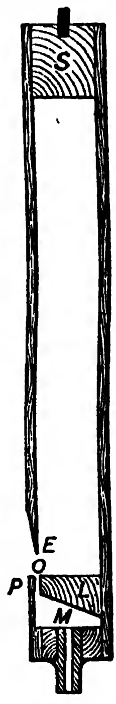


FIG. 33·14.—Organ Pipe.

Open organ pipes normally emit both the even and odd harmonics of the fundamental, whereas closed pipes only sound the fundamental and its odd harmonics. Thus an open and a closed organ pipe emitting the same fundamental differ in quality owing to the different overtones which arise in each [cf. p. 623].

The harmonics are produced in organ pipes by increasing the air blast, but organ builders have various devices for suppressing one or more of the harmonics. It is in this way that a definite quality is given to the note emitted by a pipe.

Galton's Whistle.—In its simplest form this resembles a stopped pipe. It is about 1 mm. in diameter and its length may be varied from zero to 5 cm. The whistle is blown and the frequency of the note adjusted by moving a piston which closes the tube. Notes beyond the upper limit of audibility are easily produced.

Manometric Flames.—To study the variations in pressure in an organ pipe a circular aperture is drilled at any desired point in the wall and then covered with a rubber diaphragm, A, Fig. 33·15. This membrane constitutes one side of a small chamber C into which gas is led. The gas escapes through a small orifice where it is burned. If the air pressure at A suffers a momentary change, a corresponding change takes place in the length of the jet since

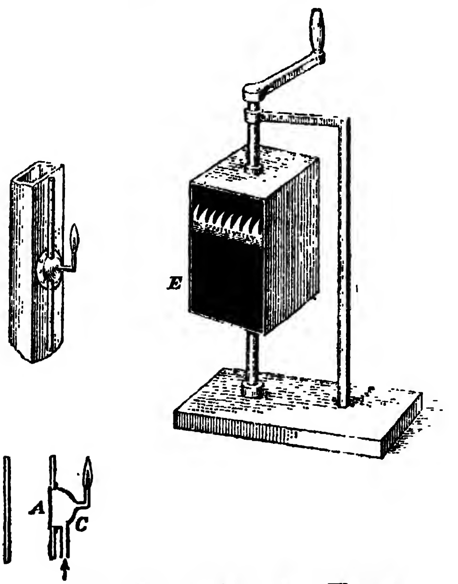


FIG. 33·15.—Manometer Flames.

the membrane moves in consequence of the pressure variation. If the changes in pressure are periodic the length of the jet also varies periodically. In general, these are too rapid to be followed with the unaided eye, but they may be made apparent by using a rotating mirror, E, Fig. 33·15. Owing to the persistence of visual impressions a number of images appear simultaneously in the mirror when it is rotated sufficiently rapidly. When the manometric flame is

at an antinode an almost continuous band of light seen in the mirror shows that the pressure variations are scarcely detectable at this point, but when the flame is at a node the upper edge of the image possesses a deeply serrated edge showing that the flame is flickering rather violently. It must be noted that the membrane responds to ***variations in pressure*** and those are ***greatest at the nodes,*** where the actual displacement is a minimum.

Rubens' Tube.—Another method of demonstrating the presence of nodes and antinodes in a vibrating column of gas is due to RUBENS. BC, Fig. 33·16, is a brass tube several metres long and about 8 cm. in diameter. Holes about 2 mm. in diameter are drilled along the top of this tube at intervals of about 2·5 cm. One end, B, of this tube is closed while a second tube about 50 cm. long slides in the

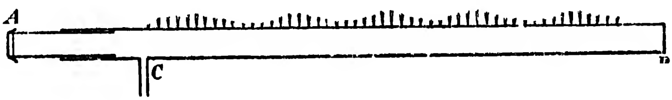


FIG. 33·16.—Rubens' Tube.

open end. A thin rubber membrane A closes the open end of the sliding tube. The side tube, C, is connected to a coal-gas supply and after a little while the gas may safely be lighted. A source of sound is placed near A and the position of the sliding tube adjusted until the gas in AB resonates: at the instant when this occurs the jets of gas vary in length as indicated in the diagram.

Kundt's Tube.—This piece of apparatus was designed for measuring the velocity of sound in solids and in gases. For this

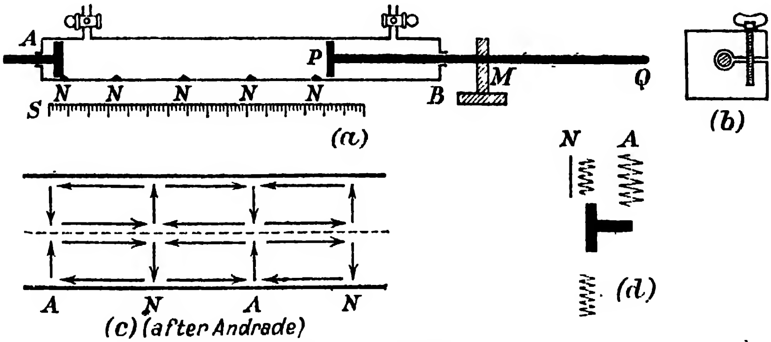


FIG. 33·17.—Kundt's Interference (or Dust) Tube.

purpose use is made of the fact that longitudinal vibrations are set up in a long rod when the latter is stroked with a resined cloth if the rod is made of metal; with a wet cloth if it is made of glass.

The apparatus consists of the rod, PQ, Fig. 33·17 (*a*), supported in a rigid stand at its centre M, the method of clamping being indicated in Fig. 33·17 (*b*). A light aluminium or cardboard disc is fastened with the aid of a drawing-pin to the end of this rod inside the tube. If the rod is of metal the attachment may be made by means of a small brass disc and a No. 6 B.A. screw. It is very important that the contact between the rod and the disc should be as perfect as possible, so that any longitudinal vibration excited in the rod shall be transmitted to the disc. To ensure this the disc may be made of copper and soldered to a short length of brass tube which just slips over the end P of the rod. The attached disc has a diameter a little less than the tube AB which contains a little recently-dried cork dust. The tube, which must be very dry, is closed by a movable piston at A. To determine the velocity of sound in the rod the latter is stroked with a resined cloth, when longitudinal vibrations, similar to those of the air in a pipe, are set up in the rod. To excite the fundamental note it is better to confine the rubbing to the rod near Q, for if it is stroked near M the overtones are often excited. The piston at A is moved until the column of air in the tube is in resonance with the rod. When this happens the cork dust is agitated somewhat violently and moved towards the nodes where it finally settles, for stationary waves have been produced in the tube. The positions of the nodes are located on a metre scale, S, at the side of the tube. The mean distance between two nodes corresponds to half a wave-length in air [cf. p. 11 for method of calculating the mean distance—if a sufficient number of nodes are formed]. Since the rod is sounding its fundamental its whole length corresponds to half a wave-length of sound travelling in it. Since the frequency, f, is the same for each motion we have, where λ_1 and λ_2 are the wave-lengths in the rod and air respectively,

$$\frac{\text{Velocity of sound in the rod}}{\text{Velocity of sound in air}} = \frac{f\lambda_1}{f\lambda_2} = \frac{\lambda_1}{\lambda_2}.$$

Hence the velocity of sound in the rod may be calculated. When this is known the value of Young's modulus for the material of the rod may be deduced from the equation

$$c = \sqrt{\frac{E}{\rho}} \qquad \text{[cf. p. 59].}$$

It cannot be expected, however, that the value for Young's modulus, determined in this way, say for brass, should be the same as that determined by stretching a wire of that material, for the vibrations in the method now under discussion take place under adiabatic conditions, whereas in the other experiment [cf. p. 140], the conditions are isothermal.

The side tubes attached to the experimental tube allow the latter to be filled with different gases when the velocity of sound in them may be determined in an analogous manner. When this information has been obtained, γ, the ratio of the two principal specific heats of the gas becomes known for

$$c = \sqrt{\frac{\gamma P}{\rho}} \qquad \text{[cf. Chap. XXXI].}$$

For monatomic gases such as argon, helium, and the vapours of mercury, sodium, and potassium, experiment shows that $\gamma = 1{\cdot}66$, a value in agreement with that deduced from the Kinetic Theory of Gases. By surrounding the tube with an electric furnace the velocity of sound in gases at high temperatures may be measured. Such studies are helpful in connection with the dissociation of gases at high temperatures for the values of γ for monatomic, diatomic, and triatomic gases are 1·66, 1·41, and 1·29 respectively. If γ is not equal to one of these values it proves that the gas under examination contains molecules having a number of atoms in them different from the number normally present.

Kundt also used such a tube at 100° C. and verified that the velocity of sound in air at that temperature was in accord with theory [cf. p. 593].

The account given above of the behaviour of Kundt's tube is by no means complete, for in addition to the tendency of the dust particles to collect at the nodes, striæ appear. An explanation of the presence of these striations has been given by König, who found it necessary to consider the viscosity of the gas in the tube. More recently Andrade has shown that the air in the tube moves in a manner similar to that depicted in Fig. 33·17 (c).

In addition, we have not yet discussed whether or not nodes exist at the 'ends of the gas column.' Now the source of the energy supplied to the gas is the vibrating rod and in order to communicate this energy to the gas in the tube, the rod must be provided with a rigidly attached piston. Now there must be a node at the stoppered end of the wave-tube as the gas molecules must be at rest there. We should expect, from analogy with the vibrations in an open tube, that the piston was at an antinode. Actually it is much nearer a node. The reason for this is that the maximum amplitude of the longitudinal vibrations in the gas will be much greater than the maximum amplitude of the vibrations at the end of the rod. At some point in between a node and antinode in the gas the maximum amplitude will be equal to that at the end of the rod. When the end of the rod occupies such a position, resonance is possible, i.e. when the rod is stroked the response of the tube is very vigorous. Fig. 33·17 (d), where

the longitudinal vibrations of the gas molecules are shown above, and those of the particles at the end of the piston, below the piston, will help to make this argument more clear.

Singing Flames.—A piece of glass tubing about 40 cm. long is heated and drawn out until the jet formed on breaking it is about 1·5 mm. in diameter. It is connected to a gas-supply and the gas lighted. When this flame, which should be about 0·5 cm. long, is inserted in a wide glass tube about a metre long, and its position gradually changed, a loud and somewhat unpleasant note is heard for a certain position of the jet—the shorter the flame, the higher the pitch of the note. On examining the flame by a rotating mirror it is found to be flickering violently. A more advanced treatment of the subject than can be given here shows that these periodic changes in the length of the flame are due to periodic supplies of heat to the air column. In consequence of these the air expands and contracts so that if the period of these is properly timed the column of air is thrown into violent and sustained vibrations.

Analytical Treatment of Stationary (or Standing) Waves.—(i) ***Reflexion at a free end, i.e. reflexion without change of phase.***—Suppose that

$$\xi_1 = \acute{a} \cos (\omega t - \eta x)$$

gives the displacement at a point A at distance x along a cord of length l at time t. This point is at a distance $(l - x)$ from the free end. Now the disturbance at x due to the reflected wave is the same as that at a point distance $2(l - x)$ from A would be if the cord were unlimited. This point is at a distance $2l - x$ from the origin. The displacement due to the reflected wave is therefore

$$\xi_2 = \acute{a} \cos [\omega t - \eta(2l - x)].$$

The resultant displacement due to the incident and reflected waves is therefore

$$\begin{aligned} \xi = \xi_1 + \xi_2 &= \acute{a} \cos (\omega t - \eta x) + \acute{a} \cos [\omega t - \eta(2l - x)], \\ &= 2\acute{a} \cos (\omega t - \eta l) \cos \eta(l - x). \end{aligned}$$

The factor $\cos \eta(l - x)$ depends, for a given cord, only on x, the position of the point considered. When it is zero, the resultant displacement is zero. The necessary condition is that

$$\eta(l - x) = (2n + 1)\frac{\pi}{2},$$

or

$$(l - x) = (2n + 1)\frac{\lambda}{4}, \text{ since } \eta = \frac{2\pi}{\lambda}.$$

$$\therefore\ x = l - \tfrac{1}{4}(2n + 1)\lambda$$

where n is a positive integer including zero.

The nodes, i.e. the points where the amplitude is zero at all times, are therefore given by

$$\begin{aligned} x &= l - \tfrac{1}{4}\lambda, && \text{when } n = 0, \\ x &= l - \tfrac{3}{4}\lambda, && \text{when } n = 1, \\ x &= l - \tfrac{5}{4}\lambda, && \text{when } n = 2, \end{aligned}$$

etc.

The amplitude of the resultant motion is $2\acute{a}\cos(\omega t - \eta l)$; a quantity which is not constant—it has a maximum value $2\acute{a}$.

(ii) ***Reflexion at a fixed end, i.e. reflexion with change of phase.*** To examine the effect of this change in phase we must increase the distance $(2l - x)$ used above by $\frac{\lambda}{2}$. The resultant displacement is therefore given by

$$\xi = \acute{a}\cos(\omega t - \eta x) + \acute{a}\cos\left[\omega t - \eta\left(2l - x + \frac{\lambda}{2}\right)\right]$$

$$= 2\acute{a}\cos\left[\omega t - \eta l - \frac{\eta\lambda}{4}\right]\cos\left[\eta(l - x) + \frac{\eta\lambda}{4}\right].$$

The condition for the cosine factor to vanish is

$$\eta\left[(l - x) + \frac{\lambda}{4}\right] = (2n + 1)\frac{\pi}{2},\ [n = 0, 1, 2, \ldots],$$

i.e.
$$x = l - \frac{1}{2}n\lambda.$$

When $n = 0$, $x = l$, i.e. the fixed end is a node.

„ $n = 1$, $x = l - \frac{\lambda}{2}$,

„ $n = 2$, $x = l - \lambda$, . . .

Thus the positions of the nodes are again determined: the antinodes occupy positions half-way between the nodes.

The Main Features of Standing Waves.—(i) Each particle executes a simple harmonic motion, except that the particles at the nodes remain fixed.

(ii) The arrangement of the particles (we still think of a cord) at any instant is that of a sine curve but twice in each period this curve becomes a straight line, i.e. the amplitude is zero at all points.

(iii) The wave-form does not advance: it only shrinks to a straight line all ordinates diminishing simultaneously. It then expands, all ordinates being reversed in sign and enlarging simultaneously. These processes continue.

THE MEASUREMENT OF FREQUENCY [1]

The Siren.—This method of finding the frequency of a fork is due to CAGNIARD DE LA TOUR. Air under pressure is forced into a cylindrical wind chest, Fig. 33·18, from which it escapes through a circular row of equidistant holes drilled in its upper surface. Above this cylinder is a movable disc having the same arrangement of holes in it. In the more simple types of this apparatus the two sets of holes are inclined as shown at (*b*). The air escaping from the stationary holes sets the disc in rotation and each time the orifices come opposite each other jets of air

[1] cf. also pp. 621.

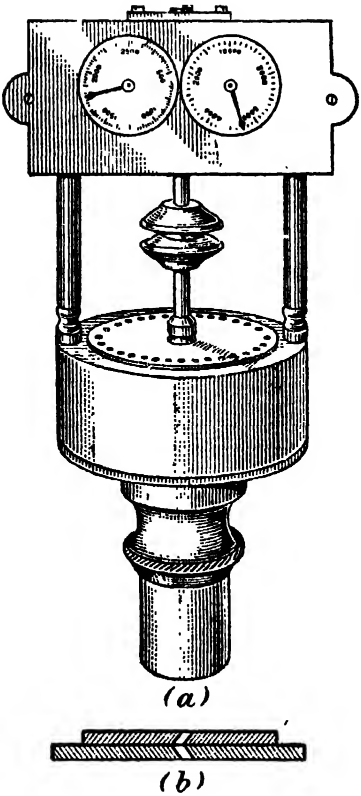


FIG. 33·18.—A Siren.

escape and compressions are produced. Suppose that there is one hole in the fixed disc of the cylindrical air chamber and that the upper and movable disc has n equidistant apertures in it. Then n 'puffs' of air escape for each revolution of the movable disc. Hence if this disc makes N revolutions per second the number of puffs per second will be Nn. In practice the fixed disc also has n equidistant apertures, but this only increases the magnitude of the disturbance in the air, i.e. the loudness of the sound. The frequency, f, of the note from the siren is given by $f = \mathrm{N}n$. The value of N is found from the speed-counter connected to the axle through a worm gear. To determine the frequency of a fork the air pressure is adjusted until the notes from the fork and siren are in unison. In deducing the speed of revolution, observations of the reading on the counter should be made at intervals of 15 seconds and the mean speed calculated by subtracting the first reading from the sixth (say), the second from the seventh, etc., as explained on p. 11.

Determinations of frequency with the above siren are not very accurate since it is difficult to keep the wind pressure constant. In the more modern forms of this instrument the apertures are vertical and the disc is driven by a motor when its speed is independent of the air pressure in the cylinder.

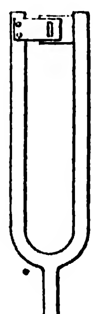


FIG. 33·19.—Tuning Fork with Slits Attached.

Stroboscopic Method.—Light metal plates are attached to the prongs of an electrically maintained tuning fork as in Fig. 33·19. Each plate has a narrow rectangular slit. When the fork is not vibrating the two slits are opposite each other so that a beam of light can pass through them. When the fork is sounding it is only when the prongs are in their mean posi-

tion that the light passes through. This occurs twice during each complete vibration of the fork. A well-illuminated white disc having a circular row of black dots is viewed through the slits. The disc is driven at a uniform speed determined by a counter attached to its axle. When both the prongs and disc are moving, the dots will, in general, appear to move. The speed of the disc is adjusted from zero until in the interval while the light is cut off one dot moves into the position just previously occupied by its predecessor.

During each complete vibration of the fork an observer obtains two views through the slots. If f is the frequency of the fork he will therefore obtain $2f$ views per second. If N is the number of revolutions made by the disc per second, and m the number of dots thereon, then Nm dots pass across any line of sight per second. But in the interval of time between two views one dot moves into the position just previously occupied by its predecessor. Hence

$$2f = \mathrm{N}m,$$

so that n may be calculated.

Stationary patterns will also be obtained when the disc makes 2N, 3N, . . . revolutions per second.

Strictly speaking, the value of the frequency obtained by this method is not the absolute frequency of the fork, for the latter carries metal pieces. To determine the absolute frequency, two nearly identical forks are necessary; by counting the number of beats occurring per second when the second fork is sounded together with the fork under investigation, (a) when the latter is loaded and (b) when it is not loaded, its absolute frequency may be deduced.

In this experiment one must be quite certain that when the dots appear stationary they are at the same distance apart as when the disc is at rest, since if the speed of the disc is only one-half the correct value, a stationary pattern with twice the number of dots appears—this is due to persistence of vision. Moreover, as the disc increases in speed, a stationary pattern is formed when its speed is two, three, etc. times too fast. By observing, in turn, the speeds of the disc, when stationary patterns are produced, a mean value for the frequency may be calculated.

The Phonic Wheel.—This device for the accurate determination of frequency is due to Rayleigh. It consists of an iron wheel, about 3 inches in diameter, having equidistant studs or cogs on its periphery—see Fig. 33·20. It is capable of revolution about a horizontal axis. A second wheel attached to the same axis helps to increase the inertia of the system. Two electromagnets are placed as shown so that the cogs almost touch the cores of the magnets, N and M; they are excited by the intermittent current

from an electrically maintained fork. The phonic wheel is caused to rotate by hand. At a certain speed the wheel will continue to

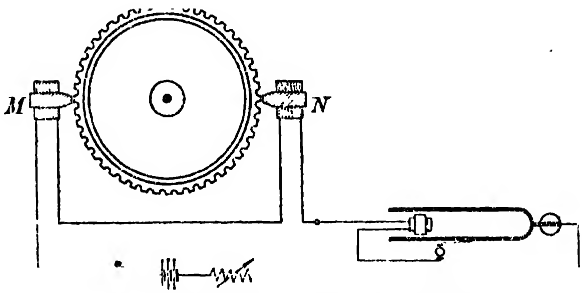


FIG. 33·20.—Rayleigh's Phonic Wheel.

run and by counting the revolutions made under these conditions the frequency of the fork may be determined. The reason for the continued motion is that when the frequency of excitation of the magnets is equal to the number of cogs passing per second, then as each cog is coming before the magnet it will be attracted and the motion persist. The motion of the wheel can be maintained for one hour so that if the time is measured accurately to one second the error should not exceed one part in three thousand.

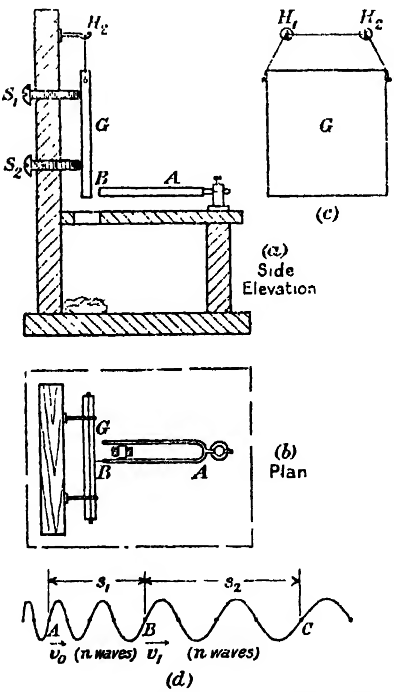


FIG. 33·21.—Falling Plate Apparatus for Determining the Frequency of a Tuning-Fork.

Determination of the Pitch of a Tuning-Fork by the Falling Plate Method.—We shall suppose that an electrically maintained tuning-fork A, Fig. 33·21 (*a*) and (*b*), is available. This is mounted in a horizontal position as indicated in the diagram. A light style, B, consisting of a pig's bristle, is attached to one prong of the fork so that about 3 mm. of the bristle project beyond the edge of the fork. Such a style is light, and yet although it yields easily to a force

at right angles to its length it returns to its zero position when that force is removed. The attachment of the hair to the fork is made with a small amount of soft wax. The end of the bristle is in contact with the smoked surface of a glass plate G. This plate is supported by a piece of cotton attached to its sides [a suitable brass holder is provided for this purpose] and passing over two hooks, H_1, H_2; Fig. 33·21 (*c*), gives a front view of the plate and its supports. S_1 and S_2 are two screws fixed in the stand carrying the apparatus. They are adjusted so that their ends are in contact with the back surface of G, and the end of S_1 is a very short distance in front of a vertical plane passing through the end of S_2 and parallel to the plate. In this way, when the plate falls, its smoked surface is made to remain in contact with the extremity of the bristle. A duster placed on the base of the stand arrests the fall of the plate.

The tuning-fork is excited and the cotton supporting the plate burnt. A wavy trace appears on the plate, and from this trace the frequency of the fork may be deduced. An example of such a trace is given in Fig. 33·21 (*d*). If the initial part of the curve is not very distinct it may be neglected by proceeding as follows. Imagine a straight line drawn down the centre of the trace and let A, B, and C be three points at which the wavy line is intersected by the straight line, and such that the same number of complete vibrations has been made in the two intervals AB and BC. Let this number be n. If the velocity of the plate at A was v_0 and the time required to make n complete waves t, then

$$s_1 = v_0t + \tfrac{1}{2}gt^2,$$

where s_1 is the distance AB, and g is the intensity of gravity. Similarly,

$$s_2 = v_1t + \tfrac{1}{2}gt^2,$$

where s_2 is the distance BC, and v_1 is the velocity of the plate at B, viz. $v_0 + gt$.

$$\text{Hence} \quad s_2 - s_1 = gt^2, \quad \text{or} \quad t = \sqrt{\frac{s_2 - s_1}{g}}.$$

The frequency of the fork, f, i.e. the number of complete vibrations it makes per second is $n \div t$,

$$\text{i.e.} \quad f = n\sqrt{\frac{g}{s_2 - s_1}}.$$

It must be remembered that this experiment determines the frequency of the fork when it is loaded with the wax and style. The method of obtaining the correction on this account is explained in connexion with the stroboscopic disc.

Lissajous' Figures.—An optical method of examining the accuracy of tuning of some interval (unison, octave, etc.) between two forks requires the apparatus shown in Fig. 33·22 (*a*). The two forks A and B are arranged so that their prongs are mutually at right angles. M_1 and M_2 are very small plane mirrors attached to the ends of the prongs of the forks nearest together. O is a small circular aperture illuminated by an electric lamp. C is a converging lens so arranged that the light from O, after falling on M_1 is reflected to M_2, and finally forms an image on a screen, S. Suppose that axes parallel to the directions of the prongs are constructed on the screen—cf. Fig. 33·22 (*b*). Let Ox be parallel to B and Oy parallel to A. Let the spot of light be brought to a focus at the origin of above axes when both forks are silent. If A alone vibrates the image will be drawn out into a straight

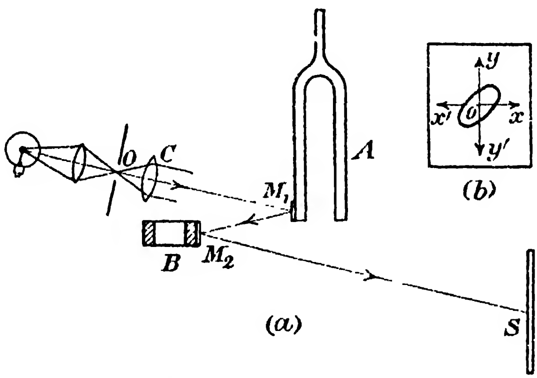


FIG. 33·22.—Lissajous' Figures: determination of $\frac{f_1}{f_2}$.

line along yOy'; if B vibrates by itself the image is a short line xOx'.

When the two forks vibrate together and they are in unison, a stationary pattern with its centre at O, Fig. 33·22 (*b*), is formed if the amplitudes of the forks remain constant. This figure will be an ellipse, circle, or straight line, depending on the phase difference of the two motions. If the unison is not exact the pattern slowly changes from one of the above three types to the others and finally regains its original shape. Suppose that this occurs in t seconds. Then in this time one fork has made one more vibration than the other. Let the frequencies of the two forks be f_1 and f_2, where $f_1 > f_2$. Then

$$f_1 t = f_2 t + 1,$$

or

$$\frac{f_1}{f_2} = 1 + \frac{1}{f_2 t}.$$

When the interval between the forks is an octave, the pattern produced is not so simple, but a cycle of changes occurs and the ratio of the frequencies may be found as above.

This method is applicable when the above ratio differs from unity by 1 part in 10^4, but with such forks it is essential that no mirrors should be attached to them—the polished sides of the prongs may be used as reflectors.

Supersonics or High-Frequency Sound Waves.—Supersonic waves are exceedingly short waves of sound the frequency being so high that they are a long way beyond the upper limit of audibility. Such waves possess some remarkable properties. The method of producing supersonics was originally developed by LANGEVIN in 1917. The work was undertaken with a view to detecting the presence of submarines by the echo of a narrow beam of high-frequency sound waves from them. Before discussing some of the properties of such waves, let us see how they may be produced.

The Piezoelectric Effect.—Quartz crystals appear in the form of hexagonal prisms with hexagonal pyramids at each end. Very often other faces are developed, but they do not concern us here. Although perfect crystals never occur, we can always imagine that such a crystal has been cut or that its outline has been drawn on a natural crystal. It must also be pointed out that any direction in a crystal parallel to a direction or axis referred to below is equivalent to the axis itself.

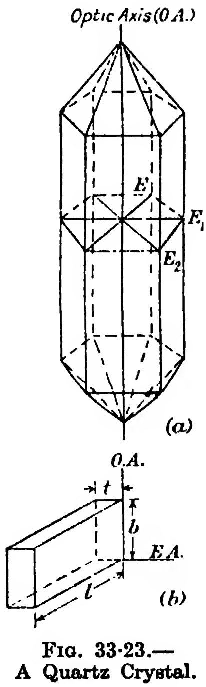


FIG. 33·23.—
A Quartz Crystal.

An ideal crystal of quartz is indicated in Fig. 33·23 (*a*). The optic axis is a straight line passing through the summits of the pyramids—or any line parallel to this. Let us imagine that a plate with its faces normal to the optic axis has been cut from the crystal. If straight lines E, E_1, and E_2, are drawn parallel to the faces of the prism, these are the electrical axes of the crystal. In Fig. 33·23 (*b*), there is shown a plate of quartz with its length, *l*, normal to one of the electric axes and to the optic axis, its breadth, *b*, parallel to the optic axis, and its thickness, *t*, parallel to the above electric axis. When such a plate is subjected to a pressure normal to its faces charges of positive and negative electricity are developed on the opposite faces. Thus there is a potential difference between the two faces of the plate. The signs of the charges are reversed when the pressure is replaced by a pull, i.e. the crystal is under tension. This phenomenon is known as the ***piezoelectric effect.***

If the faces of the quartz plate are in contact with metal sheets connected to a battery then the quartz expands or contracts by an amount depending on the strength of the field—the direction of the field determines whether or not there will be an expansion or contraction. This phenomenon is termed the ***inverse***

piezoelectric effect. Only crystals which are asymmetrical exhibit these effects.

If the applied potential difference is periodic, the quartz plate alternately contracts and expands and elastic vibrations are set up. When the frequency of the applied potential difference is equal to the natural frequency of the crystals for longitudinal vibrations in it, the amplitude of the elastic vibrations becomes very large—another example of the phenomenon known as resonance. If c is the velocity of such waves, then t, the thickness of the plate will be equal to $\frac{1}{2}\lambda$, where λ is the wave-length of the fundamental mode of vibration for the plate. The frequency, f, is therefore given by

$$= \frac{c}{\lambda} = \frac{c}{2t}.$$

The plate will also respond vigorously to applied potential differences whose frequencies are an integral multiple of f.

If the applied potential difference is V (volts), Δ, the contraction or expansion for a plate of thickness t, is given by

$$\Delta = \gamma V,$$

where γ is a constant for the given crystal. It must be noted that t does not appear explicitly in this formula—it is because the electric field is V/t and Δ/t is proportional directly to this field. For quartz,

$$\gamma = 2 \cdot 3 \times 10^{-10} \text{ cm. volt.}^{-1}$$

Hence for a p.d. of 50,000 volt.

$$\Delta = 12 \times 10^{-6} \text{ cm.}$$

Such plates are of practical importance in that they are used to stabilize the frequency of the electrical oscillations from a wireless transmitter.

High Frequency Sound Waves.—The following work was carried out by WOOD and LOOMIS in 1927, in connexion with the production of supersonics. Their apparatus is indicated in Fig. 33·24. Q is the quartz plate cut in the manner previously indicated. A and B are metal plates attached to its faces and connected to an a.c. supply of 50,000 volts and frequency 300 kilo-cycle.sec.$^{-1}$. The plate and the leads to it were immersed in oil. The frequency of the applied p.d. was adjusted until it was in resonance with the natural frequency of the plate for longitudinal waves. The oil above the plate was set vibrating. A glass plate P, 8 cm. in diameter, was then inserted in the liquid and for certain positions of this plate it experienced a considerable thrust upwards—there was then an integral number of half-waves between A and P. When one of these positions had been located the plate P could be loaded with 150 gm. and remain in position without further support.

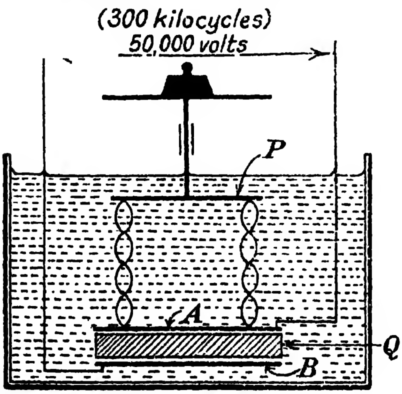


FIG. 33·24.—Supersonic Waves.

The thrust was a maximum whenever there were nodes at the lower surface of the plate for stationary waves in the oil between P and A were then formed, and the changes in pressure are greatest at the nodes [cf. p. 637].

When the plate P was removed from the oil this became heaped up to a height of 7 cm. above the rest of the oil: this protuberance was surmounted by a fountain of oil drops some of which were projected upwards to a height of 30 or 40 cm. above the oil level.

Some Experiments with Supersonic Waves.—(i) A glass tube about one metre long and 3 cm. in diameter was closed at its lower end and its inside coated with a layer of highly viscous oil. When the lower end of this tube was dipped into the vibrating oil above the plate A, rings of oil lined the tube along its whole length.

(ii) If supersonic waves are passed across the boundary formed between water and oil or mercury and water, an emulsion is formed. By means of these waves chemical reactions are accelerated and crystallization caused to begin.

(iii) A mercury thermometer was placed in the liquid above the quartz plate. It registered a temperature of 25° C. Yet the stem of the thermometer appeared to be so hot that it could no longer be held in the hand. The heat was caused by the friction between the vibrating stem and the skin of the fingers.

(iv) Supersonic waves are used for determining the depths of lakes, etc. This is derived from the time which elapses before an echo appears after a high-frequency signal has been sent downwards, and the velocity of such waves in water. This is $1{\cdot}48 \times 10^5$ cm. sec.$^{-1}$ and is independent of the frequency over a large range.

EXAMPLES XXXIII

1.—A glass tube 150 cm. long is fixed in a vertical position and filled with water which runs out slowly at the other end while a tuning-fork of frequency 495 is maintained vibrating over the upper end of the tube. At what levels of the water surface will resonance occur (*a*) if the temperature is 0° C., (*b*) if the temperature is 17° C. ? The velocity of sound in air at 0° C. may be taken as 330 metre. sec.$^{-1}$. How may the 'end correction' for such a tube be found ?

2.—Find an expression for the change in the frequency of the note heard by an observer when a source of sound is approaching him with uniform velocity. Show that the change is not quite the same if the observer moves with this same velocity towards the source when this is stationary. Account for the beats which may be heard by a stationary observer when a vibrating tuning fork is moved towards a wall.

3.—Calculate the density of the material of a sonometer wire 1 metre long and 0·70 mm. in diameter if, when stretched with a load of 20 kilograms, the first overtone it gives when vibrating transversely has a frequency of 250 cycle. sec.$^{-1}$. [Take g = 1000 cm. sec.$^{-2}$.]

4.—A sonometer is arranged to emit a note of definite frequency. How must the tension be varied to increase the frequency of the note in the ratio $\frac{5}{3}$? If the tension were maintained constant in what other way could the same change in frequency be made ?

5.—If 6 beats per second are produced by the fundamental notes of two organ pipes sounded together when the temperature is − 10° C., calculate the number of beats when the same pipes are sounded together and the temperature is 30° C.

6.—A brass rod is clamped at its middle point and stroked with a resined cloth. Describe the apparatus necessary to determine the velocity of sound in brass and show how you would deduce your result. Also describe how such an apparatus may be used to determine the ratio of the two principal specific heats of carbon-dioxide.

7.—Describe a direct method of determining the frequency of a tuning-fork. If you were provided with a tuning-fork of known frequency and another whose frequency only differed slightly from it, describe how you would determine the frequency of the second fork.

8.—An open organ pipe and a stopped organ pipe are constructed to give notes of the same pitch. Discuss the relative dimensions of the pipes, and account for the difference in quality of the two notes.

9.—Describe a method of measuring directly the frequency of vibration of a tuning-fork. A fork of unknown frequency gives 4 beats per second when sounded with another fork of frequency 256 cycle. sec.$^{-1}$. The fork is loaded with a piece of wax and it again gives 4 beats per second with the standard fork. How do you account for this result ?

10.—A pipe 160 cm. long, closed at one end, resonates to a tuning-fork of frequency 270 cycle.sec.$^{-1}$, and another pipe 252 cm. long and closed at one end resonates to a fork of frequency 240 cycle.sec.$^{-1}$. Find from these observations a value for the velocity of sound in air, at the temperature prevailing when these experimental results were obtained.

11.—A closed brass pipe emits a note of frequency 486 cycle.sec.$^{-1}$ at 20° C. If the coefficient of linear expansion of brass may be taken as $2{\cdot}0 \times 10^{-5}$ deg.$^{-1}$ C., determine a value for the frequency of the note when the pipe is sounded at 0° C. [Neglect end-corrections.]

12.—Alternating current of frequency 50 cycle.sec.$^{-1}$ flows through a stretched wire of length one metre and mass 9·81 gm. The mid-point of the wire lies between the poles of a horse-shoe magnet and the wire resonates. What is the tension in the wire ?

13.—A stiff wire is stretched by a force which increases its length by 0·1 per cent. Compare the frequencies of the fundamental notes emitted when it vibrates (*a*) longitudinally, (*b*) transversely.

14.—Explain the method of comparing the frequencies of two tuning-forks by observations on Lissajous' figures.

If in an experiment of this kind one fork is of frequency 256 cycle.sec.$^{-1}$ and a circular figure occurs every 6 sec., what deduction may be made about the frequency of the second fork ?

15.—Describe how the vibrations of a tuning-fork may be electrically maintained and explain how their frequency may be measured by a stroboscopic method.

A stroboscopic disc is revolving at the rate of 10 rev.sec.$^{-1}$ and the pattern on it consists of 30 white 'spots' arranged at regular intervals round its periphery. When examined through slits attached to a vibrating tuning-fork the spots appear to recede at the rate of one spot every five seconds. What is the frequency of the fork ?

CHAPTER XXXIV

AUDITION AND THE MUSICAL SCALE

The Anatomy of the Ear.—The structure of the organ of hearing is somewhat as follows:—It consists of an external or ***outer ear*** which is a plate of elastic cartilage covered with skin. This catches the sound waves from whence they are conducted via the ***external auditory canal*** to the ***tympanic or drum-like membrane.*** The vibrations of this membrane are communicated to a second membrane by means of a chain of ***ossicles*** or small bones. The oscillations of this membrane are communicated to a fluid contained in the canals of the temporal bone. This excites the ***sense-cells*** which in their turn affect the ***auditory nerve*** which communicates with the brain.

The factors enabling us to judge the direction whence a sound comes have not been definitely established, but it is fairly certain that an important factor is the difference in the intensity of the sound at each ear for, in general, the head will screen one ear from the oncoming waves. From such a difference previous experience alone enables us to fix the direction of the source. Some animals are capable of moving certain portions of their outer ears and the corresponding variations in intensity may enable them to fix this direction more precisely. Recent experiments have revealed the fact that difference in phase is another contributory factor, as the following experiment shows:—A long rubber tube is held in each ear and both are connected to a wider tube leading to another room where there is a source of sound. The rooms should be such that no direct sound reaches the observer. The apparent locality of the sound varies if the length of one of the rubber tubes is changed. This may be accomplished by including in one of the branches two brass tubes, one sliding easily in the other.

The Limits of Audibility.—Helmholtz, working with long tuning forks and organ pipes, found that vibrations less in number than about 30 per second failed to stimulate the auditory nerve. This represents the ***lower limit of audibility.*** By using a Galton's whistle the ***upper limit*** may be shown to be about 30,000 vibrations

per second. This, however, varies for different persons and tends to become less with advancing years.

The Analysis of a Complex Wave Motion. Fourier's Theorem.—We have already shown [cf. Chap. XXXI] that two S.H.Ms. may be compounded to produce another periodic motion having a more complex wave-form. In the same way, three or more S.H.Ms. may be compounded to produce a very complex although still periodic motion. FOURIER, in 1819, proved that any periodic motion, however complex, could be analysed into a number of S.H.Ms. the frequencies of which bore a simple relation to that of the fundamental. Now although a proof of this theorem is far too difficult for discussion here, it would at least be interesting if we could discover whether or not vibrations corresponding to these components are actually present in a wave whose form is complex. If they are, Fourier's theorem will represent physical facts and be more than a mere mathematical tool for simplifying any necessary calculations. Helmholtz found that such frequencies were actually present whenever the wave-form of the sound was complex. For these experiments he designed a special form of resonator.

Helmholtz's Resonators and Timbre.—Helmholtz found it necessary to construct this type of resonator instead of using columns of air in organ pipes, etc., because, although these latter do respond, the resonance is not sharply defined—a very necessary condition if the analysis of a sound is to be at all correct. One of his resonators is shown in Fig. 34·1. It consists of a large spherical glass vessel having a cylindrical neck, A, small in comparison with the capacity of the spherical portion of the resonator. B is a narrow stem which could be placed near to the ear. Each resonator only responds to a definite note having the same frequency as its own fundamental, the response to any other being exceptionally weak. A series of resonators were made and many musical notes produced in a variety of ways analysed by determining the resonators which responded in any given instance. It was found, for example, that the first three overtones, i.e. the 2nd, 3rd, and 4th harmonics, were present in the note from a piano and that they were fairly strong. The next three were feeble, whilst the seventh was absent. The absence of this particular overtone is necessary, for, otherwise, discord would

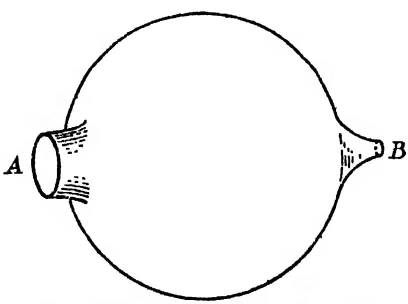


FIG. 34·1.—Helmholtz Resonator.

be present. The peculiar timbre of a violin is due to the fact that the first seven overtones are present.

In other modifications of this resonator the spherical portion is replaced by a brass cylinder and instead of using the ear to detect the response of a resonator the tube B is connected to a manometric flame which is examined by a rotating mirror. A series of such resonators are made and any flickerings of the manometric flames connected to each resonator indicate the presence of corresponding frequencies in the note examined.

Theory of a Helmholtz Resonator.—It will be assumed that the resonator consists of a globe, of volume S, to which is attached

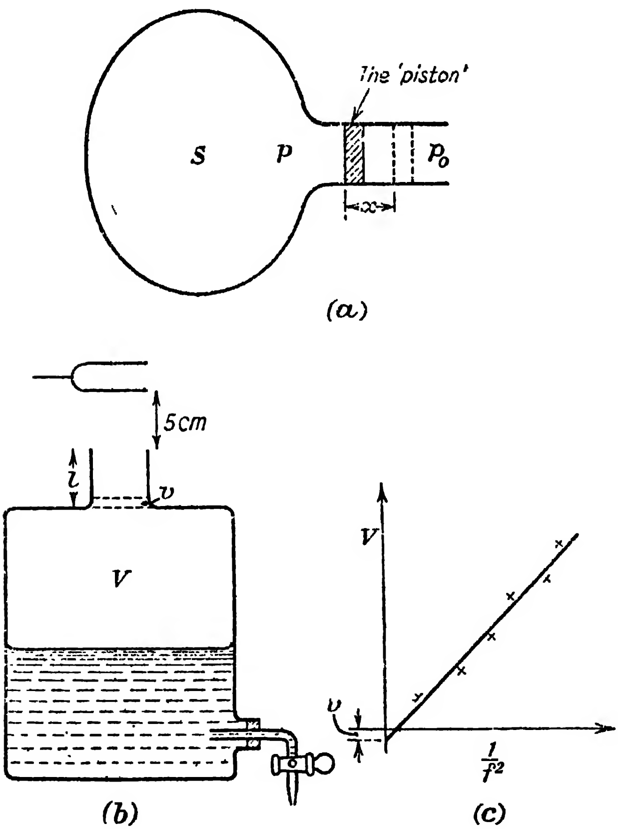


Fig. 34·2.

a fairly wide neck of cross-sectional area A—cf. Fig. 34·2 (*a*). The air in the neck is considered to act as a piston: its mass is small but the kinetic energy of the system is mainly resident in this

mass since the mean velocity of the air in the neck is much greater than that of the air in the globe. Let m be the mass of air in the neck and suppose that when the 'piston' is displaced outwards a small distance x from its equilibrium position p and p_0 are respectively the pressures inside and outside the resonator. The motion of the piston is given by

$$m\ddot{x} = (p - p_0)\mathrm{A}.$$

If it is further assumed that reversible adiabatic conditions prevail,

$$p(\mathrm{S} + \mathrm{A}x)^\gamma = p_0\mathrm{S}^\gamma,$$

where γ is the ratio of the two principal specific heats for air. This gives

$$p\left(1 + \frac{\mathrm{A}}{\mathrm{S}}x\right)^\gamma = p_0, \qquad \text{or} \qquad p - p_0 = -\gamma p\mathrm{A}x/\mathrm{S}, \text{ since } x \text{ is small.}$$

Hence the equation expressing the motion becomes

$$m\ddot{x} + \frac{\gamma p\mathrm{A}^2 x}{\mathrm{S}} = 0.$$

To solve this equation it is written

$$\ddot{x} + \omega^2 x = 0.$$

Multiplying throughout by $2\dot{x}$ and integrating we get

$$(\dot{x})^2 + \omega^2 x^2 = \omega^2 a^2,$$

since the constant of integration is $\omega^2 a^2$, because $\dot{x}$ is zero when $x = a$, say.

$$\therefore\ \dot{x} = \omega(a^2 - x^2)^{\frac{1}{2}}.$$

$$\therefore\ \sin^{-1}\frac{x}{a} = \omega t + \mathrm{C}. \qquad [\mathrm{C} \text{ is a constant.}]$$

If $t = 0$ when $x = 0$, $\mathrm{C} = 0$, and we have

$$x = a \sin \omega t.$$

Hence the frequency is f, which is $\dfrac{\omega}{2\pi}$ or $\dfrac{1}{2\pi}\sqrt{\dfrac{\gamma p\mathrm{A}^2}{\mathrm{S}m}}$. If l is the length of the neck, $m = \mathrm{A}\rho l$, and since $c = \sqrt{\dfrac{\gamma p}{\rho}}$, we have

$$f = \frac{c}{2\pi}\sqrt{\frac{\mathrm{A}}{l\mathrm{S}}},$$

i.e.

$$f \propto \mathrm{S}^{-\frac{1}{2}}.$$

Experiment with a Resonance Bottle.—A suitable 'bottle' for investigating experimentally the relation $f \propto \mathrm{S}^{-\frac{1}{2}}$ consists of a cylindrical brass drum provided with a neck as shown in Fig. 34·2 (*b*). This resonator is first filled with water to the base of the neck. A high-frequency fork is held about 5 cm. above the

neck and water is withdrawn until there is resonance between the fork and the bottle. The mass of water withdrawn (or its volume) is determined. This is repeated with several tuning forks of lower frequency.

Let V be the volume of water withdrawn when the resonator responds to a note of frequency f. Let $S = V + v$, where v is a correction term. Then according to the theory above, $f^2 \propto (V + v)^{-1}$. If therefore f^{-2} is plotted against V a straight line should be obtained—cf. Fig. 34·2 (*c*). Its intercept on the V-axis will be $- v$, but this term is usually so small that it cannot be determined accurately.

The mass of air in the neck may be altered by fixing a tight-fitting paper collar to it. Then the mass of air in the neck will be proportional to the length, l, of the neck. The above theory shows that $f \propto l^{-\frac{1}{2}}$. This may be verified. [If this part of the experiment is not attempted an ordinary reagent bottle may be used; this should be empty at first and water then added to it.]

Speech.—The vibrations of two stretched membranes situated within the larynx and termed the ***vocal cords*** are responsible for speech. They form the edges of a narrow slit and their vibrations are caused when air from the lungs is forced past them. Their tension and distance apart may be controlled at will. The pitch of a note is determined by the tension in the vocal cords, but its timbre is produced by resonance in the cavities in the throat, mouth, and nose.

The Musical Scale. Its Intervals and Notation.—Let us assume that a musical note is produced when the frequency of vibration is f_1 while another note has a frequency of f_2. Then if $f_1 > f_2$, the ratio $\frac{f_1}{f_2}$ is termed the interval between these notes. Similarly the interval between f_2 and f_3, $(f_2 > f_3)$, is the ratio $\frac{f_2}{f_3}$. Since the interval between f_1 and f_3 is $\frac{f_1}{f_3}$ it follows that the 'sum' of two intervals is equal to their product. Experience shows that the effect arising when two notes having an arbitrary interval are sounded together is not always pleasant, i.e. they are not always ***concordant,*** but ***discordant.*** For notes having frequencies p, q, and r, the effect on sounding them together is pleasant when $p : q : r = 4 : 5 : 6$.

When the interval between two notes is 2, i.e. the frequency of one is twice that of the other, that interval is termed an ***octave.*** Between a given note, say middle C, and its octave six other notes have been introduced forming a musical scale. These eight notes are indicated by the letters C, D, E, F, G, A, B, and *c*. The last

member of this octave is the first of the next one, viz. *c*, *d*, *e*, etc. For octaves higher and lower than these accents and suffixes are used. The ratios of the frequencies of the notes in the scale commonly used are shown below :—

	C	D	E	F	G	A	B	*c*
Ratios of frequency	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2
Intervals . . .	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$	

The most rational explanation of the evolution of this particular scale has been given by HELMHOLTZ. It is well known that any note produced by a musical instrument, including the human voice but not the tuning-fork, consists of a fundamental and some of its harmonics. When two notes are sounded together the effect will only be pleasing if there is concord not only between the fundamentals but also between the overtones, which may be present. This happens when the beats produced lie outside that range of frequencies which annoy the ear. The above scale was chosen so that when a melody is played the effect shall always be pleasing. Of course the notes are not all sounded together in actual practice so that beats are absent, yet, unless there is concord when they are so sounded, the transition from one note to another is too abrupt for the effect to be pleasing.

An examination of the above table shows that the following numerical relations exist between the notes in an octave :—

$$C : E : G = 4 : 5 : 6$$
$$G : B : d = 4 : 5 : 6$$
$$F : A : c = 4 : 5 : 6$$

These particular sets are called the ***harmonic triads*** : they produce a pleasant effect when sounded simultaneously. These particular triads are respectively the ***tonic, dominant*** and ***subdominant*** triads. When the members of one of these triads are sounded with another note which is an octave above the lowest member, the whole constitutes a ***major chord***.

The intervals existing between notes in the above scale are either $\frac{9}{8}$, $\frac{10}{9}$, or $\frac{16}{15}$. The first two intervals, although not exactly equal, are called a ***tone***, while the last is a ***semi-tone***. The difference between the two tones which is the quotient obtained by dividing one by the other is equal to $\frac{81}{80}$. This difference is called a ***comma***.

Musicians find that the number of notes in the above scale is not sufficient for their requirements so that extra notes, obtained by

raising or lowering the above notes by an interval equal to $\frac{25}{24}$, have been introduced. Thus A becomes A♯ [A 'sharp'] when the pitch is raised by this amount, and B, on being lowered by this same amount, becomes B♭ [B 'flat'].

Musical Temperament.—The number of notes becomes too many when a scale in strict accord with the above principles is constructed, for it must be remembered that any note in the scale may serve as the keynote from which all others may be derived. To avoid this difficulty the scale has been slightly adjusted so that a certain amount of discord is introduced. Such a scale is said to have been ***tempered.*** Several such scales having a minimum amount of discord exist, but the one in general use is the scale of ***equal temperament.*** The octaves remain as before, but eleven notes are introduced between them, each interval being $2^{\frac{1}{12}}$, i.e. 1·0595. These twelve notes constitute the ***chromatic scale*** and the intervals associated with the harmonic triads in it are 1 : 1·2599 : 1·4893, instead of 1 : 1·25 : 1·50.

The Reproduction of Sound.—One of the most successful devices for the reproduction of sound is due to EDISON. Diagrams of a phonograph and gramophone are given in Fig. 34·3. In the

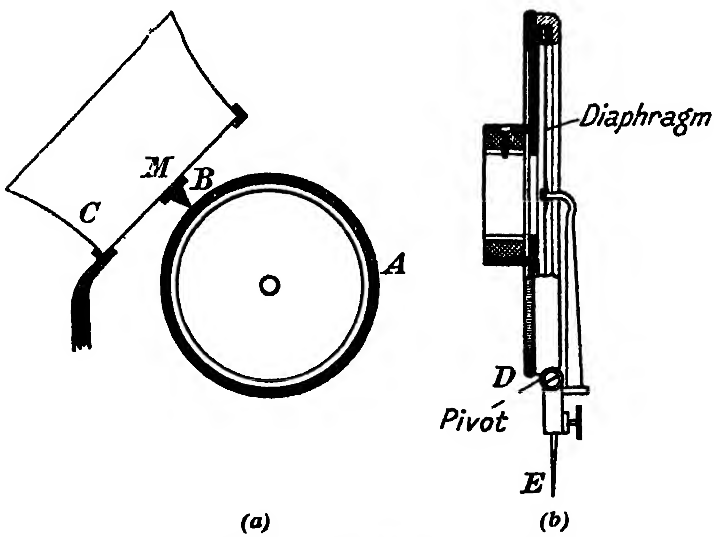


FIG. 34·3.—The Reproduction of Sound.

phonograph [Fig. 34·3 (*a*)], the earlier of the two instruments, the sound waves are caught by a cone or horn C and impinge upon the membrane M. B is a light style attached to the centre of the diaphragm and in contact with a special wax which coats the

cylinder A. This cylinder may be rotated about a horizontal axis. When it rotates and no sound-waves are incident on the membrane a groove of uniform depth is cut in the wax. But when sound waves fall upon the membrane it vibrates so that the depth of the groove varies. During this process the wax is soft, but after the impression has been taken it is allowed to harden. When the 'record' thus obtained is rotated at the same constant speed, the needle being placed at the starting-point, the end of the needle follows the groove, thus causing the membrane to vibrate and reproduce the sound. To obtain a large volume of sound the membrane is fixed at the end of a horn and the perfection of the reproduction depends upon the resonating qualities of this horn.

The diaphragm in a gramophone is attached by means of a small quantity of wax to a style moving about an axis D normal to the plane of the paper [Fig. 34·3 (*b*)]. In consequence of this the excursions of the free end, E, of the style make side cuts in an otherwise spiral groove on a wax disc in contact with it and which rotates about a vertical axis.

PART V

MAGNETISM AND ELECTRICITY

CHAPTER XXXV

ELECTROSTATICS

Introductory.—The name ***electricity*** is given to a certain invisible agent of which we are only cognizant through the effects produced by it. Although we have little idea of the true nature of electricity it is possible to give a rational explanation of these manifestations. The science of electricity, like that of magnetism, dates from the times of the ancient Greeks. This people was acquainted with the magnetic properties of lodestone; it also knew that when amber is rubbed with another substance it acquires the power of attracting small bodies to itself. These two facts, although apparently so dissimilar, are really very closely connected. It is now known that magnetism ***in motion*** produces effects similar to those due to electricity ***at rest,*** while a constant direct current of electricity [i.e. electricity ***in motion***] produces a ***stationary*** magnetic effect. Electricity is neither matter nor energy; yet it is usually associated with matter, and work must be done in transferring it from one place to another. Modern civilization owes a great debt to electricity and it is probable that this debt will increase rapidly in the future. Electricity has come to play such an important rôle because when it has been 'generated' at one station it may be transferred to another and there used in the production of heat, light, and mechanical energy. Until the last two decades of the nineteenth century material conductors were thought to be necessary, but wider knowledge has made possible wireless telegraphy where the transmitting medium appears to be space. As we find this difficult to conceive we imagine a medium filling all space—the æther—and think of it as the transmitting agency.

Electrical Attraction.—When a piece of ebonite, sealing-wax, or a glass rod is rubbed with dry flannel or silk, it acquires the property of attracting light objects, such as bits of paper, straw, etc. The Greeks discovered that amber, or, as they termed it, *ἤλεκτρον,*

behaved in this way. It was left to Dr. Gilbert (1600) to show that other bodies also acquired the same property after being similarly treated. Such bodies are said to have been ***electrified*** by friction.

Instead of using small objects to determine whether or not a body is electrified the following more sensitive apparatus may be employed. A small pith-ball is supported by a silk thread and the body under test brought near to it. If the ball is attracted, the body is electrified. This experiment is not a certain proof that the body is electrified, for if a charged piece of wax, supported in a stirrup, is similarly suspended it will be attracted when a metal rod held in the hand is brought near to it, yet the metal rod is uncharged, for it is earthed. Hence, in our first experiment the pith-ball may have been charged.

Electrical Repulsion.—If a glass rod, suspended by a silk thread, is rubbed with silk and then a second glass rod similarly treated brought near, the suspended rod will not be attracted but repelled, i.e. similarly electrified bodies repel one another. Since non-electrified bodies do not exhibit this property, repulsion is the only sure test that a body is electrified. This phenomenon of electrical repulsion explains the following facts which will have been noticed when a charged body is brought near to small objects. After such objects touch an electrified body they fall off, i.e. they are repelled. This fact was noticed by von Guericke in the seventeenth century. The reason for the above phenomenon is that after the bodies have touched the electrified body the charge on each is wholly like that residing on the charged body, so that electrical repulsion ensues.

If two uncharged pith-balls are suspended side by side and an electrified rod brought near to them, both are attracted by the rod. If they touch the rod, each acquires a charge similar to that on the rod, so that repulsion takes place. This repulsion is greatest when the rod is present although it will still persist, but in diminished amount, when the rod is removed, for the two balls have acquired similar charges. The phenomenon of repulsion is well observed when some persons brush their hair on a dry day. The hairs become charged and so repel one another.

The observations of Robert Symmer [1759] on the attractions and repulsions of charged bodies are at least amusing. He was in the habit of wearing two pairs of stockings simultaneously, a worsted pair for comfort and a silk pair for appearance. In pulling off his stockings he noticed that they gave a crackling noise, and sometimes they even emitted sparks when taken off in the dark. On taking the two stockings off together from the foot and then drawing the one from inside the other, he found that both became inflated

so as to reproduce the shape of the foot, and exhibited attractions and repulsions at a distance of as much as a foot and a half.[1]

The Detection of Electricity.—If an ebonite rod is electrified by rubbing it with silk, it possesses the power of attracting small pith-balls. When these balls touch the rod they become electrified by contact and are then thrust off from the rod. When two pith-balls are suspended by separate pieces of silk from the same point, and are electrified by contact with an ebonite rod, the two balls separate. If now a glass rod is similarly rubbed with silk, when it ***approaches*** the two balls they tend to fall together. We therefore conclude that the charge on the glass is opposite in sign to that on the pith-balls. Similar results can be obtained with a ***gold-leaf electroscope*** [cf. Fig. 35·8, p. 674]. This consists of a metallic box, C, which is preferably earthed in order to increase the sensitivity of the instrument [the reason for this will be given later—cf. p. 682]; two sides of the box are made of sheet glass for purposes of observation. Through an insulating boss, D, [made of sulphur] in the top of the box is inserted a metal rod which carries a metal disc at its top, whilst the portion inside the box is flattened out, and a piece of gold leaf attached to it. [Sometimes two leaves are used.] If a charged rod is brought near to the electroscope the leaves diverge and collapse again when the rod is removed; when a charged rod ***touches*** the metal disc or cap the leaves diverge and remain diverged when the rod is removed.

If the electroscope is charged initially, the divergence of the leaves increases when a body having a similar kind of charge is brought near: on the approach of a body with a different kind of charge the divergence decreases. [If this latter body is brought closer to the electroscope the divergence of the leaves may be reduced to zero and then increase.]

The existence of two types of electricity was first established about 1733 by du Fay, superintendent of gardens to the King of France. He found that a piece of gold leaf, electrified by contact with a piece of excited glass, was attracted when brought close to a piece of resin which had been electrified. Since both the gold leaf and resin were electrified, du Fay expected to observe the repulsion of the two bodies. From further experiments it was concluded that there were two types of electricity—one similar to that found on glass when rubbed by silk, the other to that on ebonite rubbed with fur. These are now termed ***positive*** and ***negative electricity*** respectively.

Insulators and Conductors.—For many years it was believed that only non-metallic bodies were susceptible to electrification,

[1] Cf. Jeans, *Magnetism and Electricity*, p. 11 (Cambridge University Press).

but this idea was corrected when STEPHEN GRAY about 1730 discovered that bodies could be divided into two classes, namely those through which electricity will pass (conductors) and those which prevent its passage (insulators). If a metallic tube is attached to a glass rod, the glass rod attracts small pith-balls after the whole has been rubbed, the metal being in the hand [i.e. earthed]. The metal part does not display this phenomenon. When, however, the rubbing is repeated with the glass held in the hand, then the metal retains its state of electrification—it is the glass which prevents the charge from escaping.

The above experiment shows that substances may be divided roughly into two classes—***insulators,*** which retain their charge on being excited electrically, and ***conductors,*** which lose their charge if they are earthed. Good insulators are poor conductors of electricity and *vice versa*. Amber, bakelite, ebonite [when highly polished] and dry gases are examples of good insulators, while metals and aqueous solutions of salts and inorganic acids are good conductors. The charge on an electrified body may also be removed by passing the body through a flame, or exposing it to X-rays or radium. The terms conductor and insulator are relative ones only, for pure water is an insulator for small voltages, and yet if an insulator is wet its charge of electricity is rapidly lost. It is, therefore, better to speak of good and bad conductors of electricity rather than to use the terms conductor and insulator, and to state the conditions under which the substance considered is to be used.

The fact that dry gases are bad conductors of electricity has probably been a great blessing to the human race, for, had they been good conductors, the phenomenon of electricity might have remained undetected and unsuspected.

The 'Colour' Test for Electricity.—When a mixture of sulphur and red lead, Pb_3O_4, is dried in a desiccator, and afterwards shaken, the sulphur becomes charged negatively, whilst the red lead acquires a positive charge. The mixture, as a whole, has a zero charge, a fact which can be demonstrated by placing it inside a metal cylinder which stands on a gold-leaf electroscope. If now a charged body has the powder sprinkled over it and the body is gently tapped, the sulphur [−] adheres to it, if it is positively charged, whilst the red lead [+] adheres to it, if it is negatively charged.

Quantity of Electricity.—If a tin or metal cylinder is placed upon the disc of an electroscope and a charged body is placed ***inside*** the cylinder, the leaves diverge and the divergence is constant irrespective of the position of the charged body, providing that it is well within the tin. If the charged body is removed and replaced, the divergence is the same. This constancy is attributed

to the fact that there is a definite *quantity* of electricity associated with the charged body.

The Torsion Balance.—The torsion balance was used by COULOMB for the purpose of measuring the force of repulsion between similarly electrified spheres by balancing the moment of this force about a definite point against the couple exerted by a wire when the latter is strained by twisting it from its position of rest. One form of this instrument is indicated in Fig. 35·1. It consists of a light lever suspended by a fine silver wire within a cylindrical glass case. One end of the lever carries a small spherical pith-ball, A, covered with gilt, so that any charge given to it is distributed uniformly over its surface. The lever is suspended so that it rests in a horizontal position. The silver wire is about 2 feet long and its upper end is attached to a brass head which may be rotated about a vertical axis. A measure of this rotation is given by a pointer rigidly fixed to the brass head and moving over a circular scale graduated in degrees. An insulated second pith-ball, B, may be introduced through an aperture in the cover of the instrument: it is supported in the same horizontal plane as A. To keep the inside of the apparatus dry and thereby improve the insulation, a small vessel containing pumice soaked in sulphuric acid is placed in the bottom of the case.

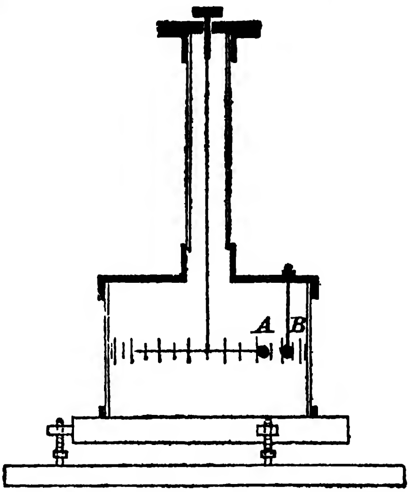


FIG. 35·1.—Coulomb's Torsion Balance.

To measure the force of repulsion between two like charges the following method is adopted:—The position of the torsion head is adjusted until the two balls A and B are in contact. The ball B is then removed and charged. When it is replaced the charge is shared by the two spheres, the charges on each then being identical since the two spheres are equal. In consequence of the like charges on the spheres they are repelled, but only A moves since the other is fixed. This produces a twist in the wire. The magnitude of the repelling force decreases as the distance between the charges increases, but the restoring couple due to the torsion in the wire increases under the same conditions. Ultimately a position of equilibrium is attained in which the moment of the repelling force about the axis of suspension is balanced by the couple arising

from the twist in the wire. Experiment shows that the couple due to torsion is proportional to the angle through which one end of the suspension is turned relatively to the other.

The Law of Force between Charged Particles.—The force between two charged bodies in air, whose dimensions[1] are small compared with their distance apart, is directly proportional to the product of their charges, and inversely proportional to the square of their distance apart, the force being one of repulsion (positive) or one of attraction (negative) according as the two charges are of the same or of opposite kinds.

To Verify the Inverse Square Law by Coulomb's Method.—When the distance between the balls A and B is small it may be assumed that the distance between them is halved when the angle they subtend at O is reduced to half its original value. Let us suppose that when B was charged and placed in position that A was repelled through an angle of 34°: this was also a measure of the twist in the wire which balanced the repelling force between the two spheres. To reduce the angular deflexion between the spheres to 17° it was necessary to rotate the torsion head through 119° *in the opposite direction* so that the relative twist between the two ends of the torsion wire was $(119° + 17°) = 136°$. Since the distance between the spheres had been halved, the force of repulsion between them had been increased four times. These numbers verify the inverse square law.

To Verify that the Force is Proportional to the Product of the Charges.—Let us assume that when the two balls A and B had equal charges that their angular separation was θ. When the ball B was removed and allowed to share its charge with another ball equal in size to itself its charge was reduced to one-half its initial value. On replacing B in position it was found that the deflexion was less than before, and in order to increase the separation to θ it was necessary to rotate the torsion head through an angle ϕ *in the same direction* so that the relative twist between the ends of the suspension was $(\theta - \phi)$. Experiment showed that $(\theta - \phi)$ equalled $\frac{1}{2}\theta$, so that the repelling force was halved when the charge on one of the spheres was halved.

The Electrostatic Unit of Electric Quantity.—By means of a torsion balance, it has been shown that the force of repulsion between two *like* charges[2] q_1 and q_2 at distance r apart and in air, is given by the equation

$$F = \lambda \frac{q_1 q_2}{r^2}$$

[1] Such a charge is often termed a 'point charge.'

[2] Strictly speaking, these should be point charges, i.e. the charges should reside on bodies whose dimensions are small compared with their distance

where λ is a constant. This equation may be simplified by a proper choice of units which will make $\lambda = 1$. This is done by choosing our unit of electric quantity so that when $q_1 = q_2 = 1$, and $r = 1$ cm., F is equal to one dyne, for then the above equation becomes

$$1 = \lambda \, . \, \frac{1 \times 1}{1^2}, \text{ or } \lambda = 1.$$

Definition.—The unit of electric charge is that point charge, of the type of electricity found on glass when this is rubbed with silk, which, when placed one centimetre away from an equal charge in air [or better, in a vacuum], repels it with a force of one dyne.

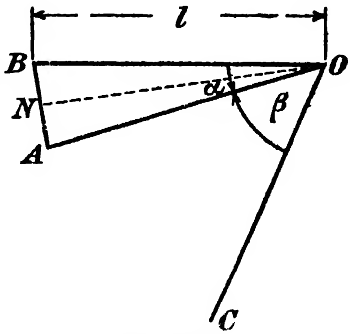


FIG. 35·2.

More Exact Theory of the Torsion Balance.—Let the charges on A and B be q so that repulsion ensues and the wire is twisted and suppose that the torsion head is rotated through an angle β ***in the opposite direction*** to reduce the angular separation to α [see Fig. 35·2]. Then $(\alpha + \beta)$ is the relative twist between the two ends of the wire and this is proportional to the repelling force F which exists when the balls occupy these particular positions, i.e.

$$\text{F.ON} = \kappa(\alpha + \beta)$$

where ON is the perpendicular from O on AB, and κ is a constant. Since ON $= l \cos \frac{\alpha}{2}$, the above equation becomes

$$\text{F}.l \cos \frac{\alpha}{2} = \kappa(\alpha + \beta).$$

If r = AB, then

$$\text{F}.r^2 = \frac{\kappa(\alpha + \beta)}{l \cos \frac{\alpha}{2}}.r^2 = \frac{\kappa(\alpha + \beta)}{l \cos \frac{\alpha}{2}}.\left(2l \sin \frac{\alpha}{2}\right)^2$$

$$= 4\kappa l(\alpha + \beta) \sin \frac{\alpha}{2} \tan \frac{\alpha}{2}.$$

To Verify the Inverse Square Law.—If F is proportional to r^{-2} the product Fr^2 should be invariable when the distance r is varied provided that the charges on the spheres remain constant: in other words $(\alpha + \beta) \sin \frac{\alpha}{2} \tan \frac{\alpha}{2}$ should remain invariable, since κ and l are constants. The necessary observations are therefore corresponding

apart. The charges we have used have been on spheres because it can be shown that the effects due to such are the same as those arising from similar charges placed at the centres of the spheres.

values α_1, β_1, α_2, β_2, etc., and if $(\alpha + \beta)\sin\frac{\alpha}{2}\tan\frac{\alpha}{2}$ is found to be constant the law will have been verified.

To Compare Charges by Means of the Torsion Balance.—Since, with the usual notation, $F = \frac{q^2}{r^2}$ [q being the charge on each sphere], the equation established above may be written

$$q^2 = 4\kappa l(\alpha + \beta)\sin\frac{\alpha}{2}\tan\frac{\alpha}{2}.$$

If q_1 and q_2 are the charges to be compared, let us assume that when q_1 is shared between A and B so that the charge on each is $\frac{q_1}{2}$, the torsion head must be rotated through β_1 to reduce the deflexion to α. Similarly when q_2 is shared between A and B, these having been discharged after the first part of the experiment, let β_2 be the angle of rotation of the head to reduce the angular separation between the spheres again to α. Then

$$\frac{\left(\frac{q_1}{2}\right)^2}{\left(\frac{q_2}{2}\right)^2} = \frac{4\kappa l(\alpha + \beta_1)\sin\frac{\alpha}{2}\tan\frac{\alpha}{2}}{4\kappa l(\alpha + \beta_2)\sin\frac{\alpha}{2}\tan\frac{\alpha}{2}}$$

i.e.
$$\frac{q_1}{q_2} = \sqrt{\frac{\alpha + \beta_1}{\alpha + \beta_2}}.$$

Example.—Two small spheres each having a mass m gm. and charge q, are suspended from a point by insulating threads, each l cm. long but of negligible mass. If θ is the angle each string makes with the vertical when equilibrium has been attained, show that

$$4mgl^2\sin^2\theta\tan\theta = q^2.$$

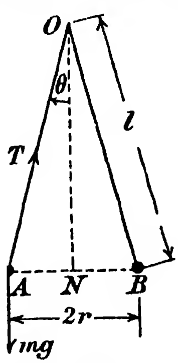


FIG. 35·3.

Let O, Fig. 35·3, be the point of suspension, while A and B are the two charged spheres. Let AB = $2r$. Consider the sphere A. It is acted upon by three forces, viz., T, the tension in the string, its weight mg acting vertically downwards, and a repelling force, F, acting in the direction BA, due to the charges on the spheres. Its magnitude is $\frac{q^2}{(2r)^2}$. Draw ON perpendicular to AB and take moments of forces about O. Then F . ON = mg . AN,

i.e.
$$\frac{q^2}{4r^2}\,.\,l\cos\theta = mg\,.\,l\sin\theta.$$

Hence
$$q^2 = 4mgl^2\sin^2\theta\tan\theta.$$

If θ is small, as is usually the case, this may be written $q^2 = 4mgl^2\theta^3$; θ is then easily deduced.

The Electric Field.—The properties of the space round a given body become modified when that body acquires an electric charge, for if other charges are now introduced into that space they experience forces, whereas such forces were absent when the body was uncharged. The space round a charged body in which these forces arise is termed an ***electric field.*** If the charged body is situated in an unlimited medium it is clear that the extent of the field increases as the sensitivity of the devices used for detecting the forces increases.

Electric Field Strength or Electric Intensity.—The strength of an electric field at a given point in air is defined, numerically, as ***the force which would be exerted on a unit positive charge placed at that point, provided that the configuration of the field were not altered by the introduction of the unit charge.*** The direction and sense of the field strength are identical with those of the above force. Hence the electric intensity at a distance r from a point charge q in air is given by $E = \frac{q}{r^2}$.

More exactly, the electric field strength is defined by the equation

$$E = \lim_{\delta q \to 0} \frac{\delta F}{\delta q} = \frac{dF}{dq},$$

where δF is the small force experienced by a small positive charge δq introduced into the field at the point where the field strength is required.

Thus, if q is the point charge to which the field is due,

$$\delta F = \frac{q \cdot \delta q}{r^2}, \text{ or } \frac{\delta F}{\delta q} = \frac{q}{r^2}.$$

Since $\frac{q}{r^2}$ is also the limiting value of $\frac{\delta F}{\delta q}$, it is the electric field strength required.

Lines and Tubes of Force.—In consequence of the electric intensity existing at all points in an electric field, it follows that a small, free, positive charge will be urged in a definite direction if placed at any point in the field : in fact, it will begin to move along the direction in which the field strength at the point considered acts. If the small charge could move without acquiring an appreciable velocity it would travel along a ***line of force,*** this being a curve such that the tangent at any point gives the direction of the electric intensity at that point. The direction in which the small positive charge tends to move is termed the positive direction of the line of force ; since repulsion takes place between like charges it follows that lines of force must have their origin on positive charges and terminate on negative ones.

The lines of force from an isolated positive point charge are

straight lines radiating outwards from that point: if the charge is negative the diagram is the same, but the positive direction of the lines is reversed.

Fig. 35·4, (*a*), (*b*), and (*c*) depict the lines of force due to equal like charges, equal unlike charges, and two like charges q and $4q$. In (*a*) there is a neutral point half-way between the charges: in

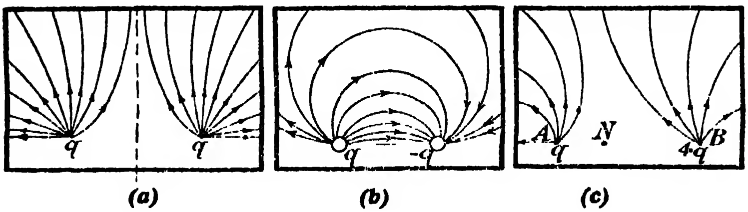


FIG. 35·4.—Lines of Electric Force.

(*c*) there is a neutral point the position of which may be determined as follows:—If N is the neutral point, i.e. the resultant field strength is zero at N, we have

$$\frac{q}{AN^2} - \frac{4q}{BN^2} = 0.$$

Hence $BN = 2 \,.\, AN$.

Such diagrams as these are useful since they give us a picture of electric fields, but we have to remember that these diagrams are drawn in one plane whereas the electric field exists in space. A more complete representation of the field due to two charges is obtained by imagining the above diagrams to be rotated about an axis passing through the charges.

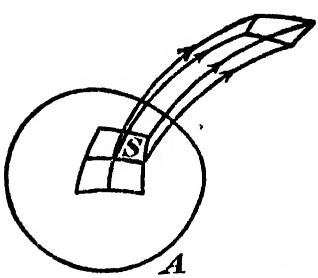


FIG. 35·5.—A Tube of Force.

Another method of depicting an electric field is as follows:—Let A, Fig. 35·5, be a positively charged body. Consider the lines of force which originate from all points on the contour of a portion S of the surface. The lines will form a tubular surface, the whole being called a ***tube of force***. If all the surface of A is divided in this manner and the corresponding tubes of force constructed they will fill the whole field and touch one another laterally. If the surface of A is divided so that each element S contains unit charge, the tubes of force arising from them are known as ***Faraday unit tubes***. Hence, if the total charge on A is q, the number of Faraday unit tubes is also q.

Electrification by Influence or Electrostatic Induction.— In the earlier part of this chapter it has been shown that bodies carrying electric charges attract or repel one another according to the signs of the charges. We also learned that attraction occurred when a charged body was brought near to one having no charge. Now it is a fundamental law in Nature that there can only be mutual action between two bodies if each is endowed with the same physical property. Thus inert matter attracts inert matter and electrically charged bodies attract or repel one another. The problem which at once presents itself therefore is to explain the electric attraction between a charged and an uncharged body. Let A, Fig. 35·6, be a positively charged sphere supported on an insulating stand, while BC is an uncharged insulated conductor. Small pith-

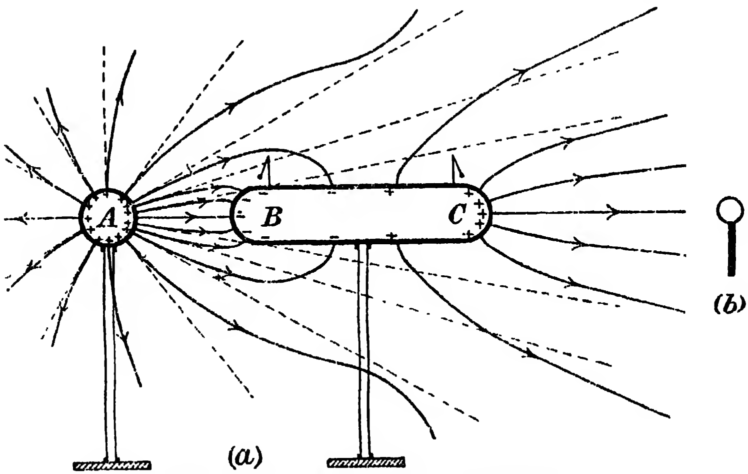


FIG. 35·6.—Electrification by Induction.

balls are placed near to the ends of BC in the way indicated. Initially these hang vertically downwards. When BC is brought near to A it will be noticed that the balls are repelled away from the surface of the conductor and that their displacements from the position of rest increase as the conductor BC is brought nearer to A. This experiment shows us that BC is charged, but it tells us nothing about the nature of the charges on it. On removing A, however, the pith-balls resume their original positions showing that BC is charged no longer. Moreover, if a pith-ball is placed about half-way between B and C, it remains undeflected during the course of the above experiment. These facts suggest that there is positive electricity on one half of the conductor and an equal amount of negative electricity on the other and that there is no electricity at the centre, but they do not indicate how it is distributed. The

sign of the electricity may be ascertained by sprinkling the surface of BC with a mixture of red lead and sulphur [cf. p. 662]. When the conductor is tapped, with a glass rod [say], the red lead adheres to the end B, while the sulphur adheres to the end C, i.e. B has acquired a negative charge, and C a positive one. If A had been charged negatively the signs of the electrification on BC would have been reversed.

This action takes place over considerable distances and even if a sheet of cardboard, glass, or ebonite is placed between A and BC. When a body becomes charged in this way it is said to have acquired its charge by ***influence*** or ***electrostatic induction.*** The phenomenon of electrostatic induction was discovered by STEPHEN GRAY in 1729.

If the conductor BC consists of two parts which are together at first but separated whilst the inducing charge is near, the two induced charges cannot neutralize each other when the inducing charge is removed, but remain on the two portions. If the inducing charge is positive, the nearer portion of this compound conductor will have a negative induced charge, while the other will have a positive one.

If the complete conductor BC is earthed while under the influence of a positive charge on A, we shall really have a compound conductor consisting of the conductor, the person touching it, and the earth. The induced positive charge will pass to the earth, so that when the finger is removed a negative charge will be found on BC even when A is no longer present.

The quantity of electricity induced on a conductor increases with the charge on the inducing body and when the distance between the two bodies is diminished. The theoretical limit would be reached when the quantity of electricity on the near end of the conductor is equal in magnitude to the charge on the inducing body, but opposite in sign, and the quantity at the far end is equal in magnitude and sign to it. In practice, however, this condition is seldom reached for when A is brought very near to the end B of the conductor, the electric intensity in the field immediately between A and B becomes so great that a minute spark passes. This is not often seen although it may be heard. After such a spark has passed and A is removed BC is found to have a positive charge since it has lost some of its negative electricity during the passage of the spark.

These experiments show that the attraction between a charged and 'an uncharged body' is really an attraction between the charge on the inducing body and the charge of opposite sign which it has induced on the nearer portion of the body which was initially without charge.

To Charge an Electroscope by Induction.—The four essential stages by which this is accomplished are shown in Fig. 35·7. A negatively charged ebonite rod, A, is brought near to an uncharged electroscope, the metallic case of which is earthed. Positive electricity is induced on the disc of the electroscope whilst negative electricity is found on the leaves. This is the reason why the

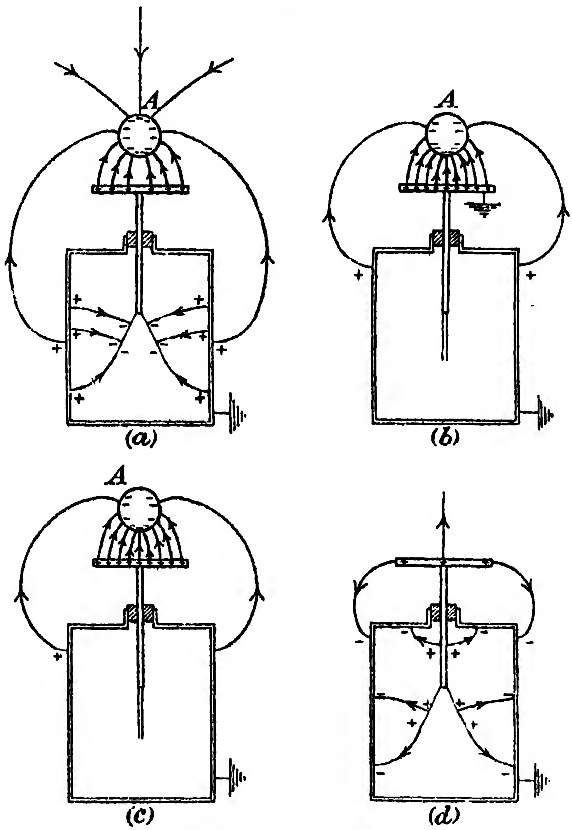


Fig. 35·7.—To Charge an Electroscope by Induction.

leaves of an electroscope diverge even when the charging rod does not touch the instrument [cf. Fig. 35·7 (*a*)]. The disc of the electroscope is then earthed—touching it with a finger will be effective [Fig. 35·7 (*b*)]. The negative charge is removed to earth but the positive charge does not so escape, for the negative charge on the rod attracts it very considerably. The finger is then removed and no change in the electroscope is observed [Fig. 35·7 (*c*)]. The rod is then removed and the positive charge on the disc spreads itself all over the surface of the disc and leaves, so

that the leaves diverge. The fact that the leaves have a positive charge can be demonstrated by bringing a positively charged rod near to the electroscope and observing the increased divergence of the leaves.

Theories of Electrification.—In the eighteenth century attempts to account for electrical phenomena usually postulated the existence of one or two imponderable fluids. FRANKLIN (1749), and others, proposed a one-fluid theory according to which all bodies in their ordinary neutral condition were assumed to possess a definite quantity of this fluid, whereas an excess or deficit of this fluid produced a positive or negative distribution respectively. Franklin further assumed that the fluid was self-repellent but attracted by ordinary matter. Hence the amount of fluid associated with a so-called neutral or uncharged body was such that the attraction between the body and the fluid in it was counterbalanced by the repulsion between the fluid in the body and the fluid external to it.

In 1759 SYMMER proposed a theory postulating the existence of two imponderable fluids. COULOMB developed this idea, maintaining that a positive state of electrification was not due to an excess of electric fluid and the negative state to a deficiency, but in the former instance the state of electrification was due to the possession of a larger portion of 'one of those active powers': in the second instance to a larger portion of the other. Moreover, a body in its natural state was unelectrified because there was 'an equal ballance of those two powers in it.'

Modern theory suggests that there is an 'atom of electricity' just as there are atoms of ordinary matter. The atom of electricity is termed an ***electron***: each electron is a definite quantity of negative electricity and its mass is $\frac{1}{1800}$-th part of that of a hydrogen atom. A negative charge is acquired when a body gains a number of electrons, whilst a deficit in the number of electrons normally present gives rise to a positive electrification. When, for example, glass is rubbed with silk, electrons are transferred from the glass to the silk so that the glass becomes charged positively and the silk negatively. The equality of the two kinds of electricity produced by friction [cf. p. 675] is at once explained and we see that there is no such thing as a generation of electricity, but that electric phenomena are due to a mere redistribution of the amount of electricity normally present in a body. Conductors permit a free passage of electrons through them whereas non-conductors only allow the electrons to suffer a small displacement from their zero positions.

The Distribution of Electricity on Bodies.—(*a*) ***Poorly Conducting Substances:*** When a charge is given to one of

these substances it is confined to that region of its surface where it has been in contact with the charging body. If the substance were a perfect insulator and there were no loss of charge through the surrounding medium, its charge would remain on its surface. In practice it is found that the charge gradually distributes itself over the body and to a less extent into its interior. Thus a piece of sealing-wax rubbed at one end only exhibits electrification at that end.

(*b*) ***Conductors:*** With conductors it is found that the charge resides wholly on their surfaces. This may be demonstrated by insulating a metal tin on a block of paraffin wax and charging the tin. The distribution of the electricity on it may be ascertained by using a ***proof plane*** [cf. Fig. 35·6 (*b*)]. This consists of a small metal disc attached to the end of a piece of sealing-wax. In use, it is brought into contact with any body under examination and then allowed to touch the cap of a gold-leaf electroscope having a charge of known sign. If the angular separation of the leaves increases, then the proof plane had a charge similar in sign to that on the electroscope. The reverse is true if the angular separation decreases. If the amount of decrease is very small the proof plane may be uncharged. To test this the electroscope should be given a charge of opposite sign to that it carried initially. If, when the plane is brought into contact with it, there is again a small decrease in the angular separation of the leaves, the proof plane will have been uncharged, but if there is a small increase in the angular separation, then the proof plane will have been carrying a small charge equal in sign to that now present on the electroscope. To verify that there was no charge on the proof plane in a suspected instance, it may be placed in contact with an uncharged electroscope. If its leaves do not diverge the plane is uncharged.

If the proof plane is allowed to touch the outside of the above tin it will always be found charged. On the other hand, if it is placed well within the tin and allowed to touch the interior, no charge will be detected on the plane when it is removed.

Also, if an insulated charged conductor is introduced into an uncharged tin supported on an insulating stand and allowed to touch it, it will be found uncharged when it is withdrawn. On testing the tin with a proof plane, however, a charge will be detected on its outside only.

Similarly, if a pair of gold leaves is supported inside a closed wire-gauze cage, the leaves do not diverge when the cage is charged.

Faraday, in order to examine this question still further, placed a charge on a butterfly net. This consisted of a conical linen-gauze bag: it was supported on an insulated ring and had silk strings attached to its apex so that it could be drawn inside out. When

examined with the aid of a proof plane and electroscope, Faraday found the charge only on the outside of the bag. When the bag was turned inside out, a charge was only detected on its outside—that portion which had previously been the interior.

Faraday's Ice-Pail Experiments.—One of a most striking series of experiments first carried out by Faraday was as follows :—A pewter ice-pail, or, as one would now use, a cylinder of perforated zinc sheet, B, Fig. 35·8, supported on an insulating stand [a block of paraffin wax] is connected by a wire to the cap of a gold-leaf electroscope C. Let us assume that a positively charged conductor A is lowered into the cylinder. As soon as the ball gets near to B the leaves of the electroscope begin to diverge owing to the inductive action of the charge, i.e. negative electricity appears on the inside of the cylinder while positive electricity is found on its outside and on the cap of the electroscope which forms part of the conductor under

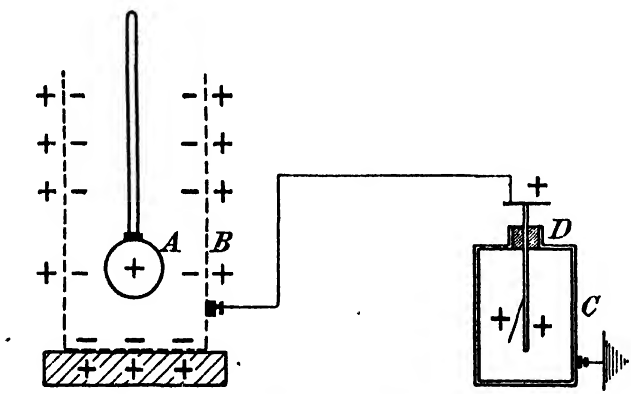


FIG. 35·8.—Faraday's Ice-Pail Experiment.

the influence of A. The divergence increases until the body A is well inside B—after this the divergence remains independent of the position of A inside B. This shows that the potential [cf. p. 677] of B and the electroscope is constant and independent of the position of A, providing the latter is well within B. If A is withdrawn, the leaves collapse. If, however, A is allowed to touch the inside of B before it is withdrawn, the divergence of the leaves is unaltered. Thus the potential of B is not changed when the positive charge on A is neutralized by the induced negative charge inside B. If now, A is removed, the divergence remains the same, i.e. there can be no charge on A after it has touched B—in other words, no charge resides on the inner surface of a hollow conductor.

The above experiment may be varied by touching the can with

one's fingers while the charged body A is inside it but before contact between A and B has occurred. The positive induced charge disappears and the leaves collapse. On removing one's fingers the leaves still remain together, a condition which is not changed when A is allowed to touch B. This is because the negative induced charge on B neutralizes the positive inducing charge. If A is withdrawn, however, after B is earthed, the leaves of the electroscope diverge, due to the fact that the induced negative charge on B distributes itself over B and C.

To Show that the Electric Field Strength inside a Charged Tin is Zero.—Let us suppose that the cylinder of the previous experiment has been charged. Let two proof planes with their metal discs touching be placed outside the cylinder so that one disc is a little farther from it than the other. If the two planes are separated while in this position and first one, and then the other, brought into contact with a charged electroscope, they will each be found charged with electricity of opposite sign. This is because they have been in an electric field and the charges induced upon them isolated by separating the component parts of the conductor while still in the field. On the other hand, however, if the two discs are placed right inside the charged cylinder, separated, withdrawn from it, and then examined, no charge will be detected on either. Hence they must have been separated in a field whose electric strength was zero.

Positive and Negative Charges produced by Friction are always Equal in Amount.—A small metal disc insulated by a wax handle is covered with a piece of ebonite, whilst another similar disc is covered with fur. The ebonite and the fur are discharged by placing them in contact with an earthed metal plate [or better, by allowing X-rays to fall on them]. The absence of electrification on them may be tested by introducing them in turn into a tin placed on top of a gold-leaf electroscope. An absence of divergence on the part of the leaves shows that they are not electrified. The two are then placed right inside the tin and rubbed together. The leaves do not diverge, showing that the total electrification on the two bodies is zero. But if either disc is withdrawn, the leaves diverge. Since both the fur and the ebonite may thus be shown to be charged while their total electrification is zero, we conclude that the positive and negative electricity are produced in equal amounts during this process.

To Show that the Surface Density of Electricity is Greatest where a Conductor is most Sharply Curved.—If a pear-shaped conductor is charged and the distribution of the charge examined with the aid of a proof plane it will be found that the density of

the surface electricity is a maximum at those points where the curvature is greatest (i.e. the radius of curvature least). For this experiment lead discs, equal in area, are bent so that each fits a particular portion of the surface to be tested. They are then mounted on sticks of sealing-wax and applied, in turn, to those portions of the conductor for which they were designed. In this way the amounts of electricity on equal areas of a charged surface may be compared by observing the divergence of the leaves of a gold-leaf electroscope when the charged discs are placed, in turn, right inside a deep tin can placed on the disc of the electroscope. If this is uncharged before each disc is introduced, the divergence is directly proportional to the charge on the disc.

EXAMPLES XXXV

1.—Define the electrostatic unit of quantity of electricity. What do you understand by the statement that equal quantities of two kinds of electricity are produced when ebonite is rubbed by fur? How would you test the accuracy of this statement?

2.—The force of attraction between two charges is 103·4 dynes when the distance apart is 6·3 cm. If one charge is numerically equal to 5 times the other, calculate the magnitude of the smaller charge.

3.—Three charges, 3, 4, and 5 positive units respectively, are placed at the corners of an equilateral triangle whose side is 10 cm. Calculate the force on the larger charge.

4.—Two charges attract one another with a force of 8·3 dynes when their distance apart is 5·2 cm. Calculate the force when the distance is trebled.

5.—Describe how you would charge an electroscope by induction.

6.—Describe a gold leaf electroscope and explain how you would use it to investigate (*a*) the distribution of charge over an insulated charged conductor, (*b*) which of the following are conductors of electricity—paper, india-rubber, chalk, a gas flame.

7.—Describe how you would show that there is no charge inside a tall insulated and charged tin and that the electric field inside is also zero.

CHAPTER XXXVI

POTENTIAL AND CAPACITANCE

Potential Difference.—When two charged conductors are connected by a wire, in general, there is a passage of electricity from one to the other : there is a redistribution of charge although the total quantity of electricity (reckoned algebraically) is constant. If, for example, the charges on the conductors are $+ 25$ and $- 8$ units respectively, the total charge before and after placing the conductors in communication is $+ 17$ units. We may examine the matter qualitatively as follows :

Suppose that two insulated metal spheres have been charged (by induction is a convenient method). If each, in turn, is placed right inside a hollow metal conductor resting on the disc of an electroscope, the leaves will diverge, the ratio of the angles through which the leaves are deflected in the two instances being a rough measure of the ratio of the charges. Now let both spheres be placed inside the above hollow conductor but without permitting them to touch. Note the deflexion of the leaves. If the spheres are then allowed to touch the divergence of the leaves is unaffected, i.e. the total charge is constant. If the spheres are tested individually afterwards, it will be found that there has usually been a redistribution of charge.

The question which arises at this stage is 'What are the conditions determining the direction in which the electricity shall flow ?' Before attempting to answer it, we must be provided with a means of obtaining definite multiples of a given amount of electricity. The following method is simple and sufficiently accurate for our present purpose. A well-insulated metal sphere is charged negatively and a small metal sphere, also insulated, placed near to it. This latter is charged by induction and if touched momentarily with the finger the positive induced charge alone remains. If the small sphere is removed and placed inside a hollow conductor the whole of its charge is given to that conductor. An equal charge may then be given to the small sphere by placing it in the same position with respect to the large one and repeating the process.

Suppose that two identical gold-leaf electroscopes are available,

and that on their discs rest two deep metal cans A and B, Fig. 36.1 (*a*). By charging the small insulated sphere already referred to, introducing it into A and allowing it to touch the sides of the can, its whole charge is given to A and the leaves of the electroscope diverge. The sphere itself is discharged and may be recharged to the same extent in the manner already indicated. Suppose that equal charges are given to A and to B. It will be found that the divergence of the leaves is different in the two instances if A and B are different in size. Now let a wire, insulated by supporting it on a stick of sealing wax, be allowed to touch both A and B simultaneously, as in Fig. 36.1(*b*). The divergences of the leaves are equal, that of the leaves on the first electroscope having increased, while that of the leaves on the second have decreased. Since there

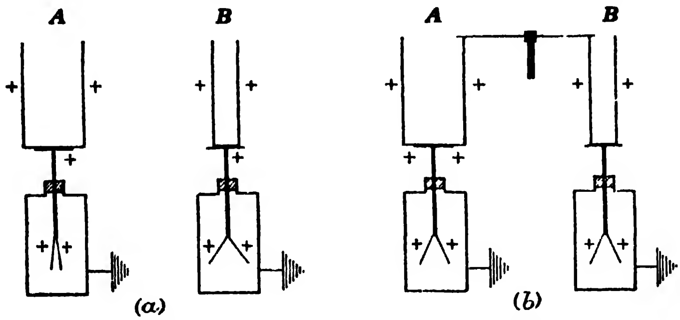


Fig. 36·1.—Introductory Experiment on Potential.

has been a change in the deflexion of the leaves it is concluded that there has been a transfer of electricity from one system to the other. Since equal quantities of electricity were given to them originally, it follows that quantity of electricity is not the factor determining the flow of electricity from one to the other. The factor which does determine it is termed ***electric potential difference*** and if two charged conductors are connected metallically one with the other, electricity flows from the conductor at the higher potential to the other. Since only potential differences may be detected, it is convenient to have a standard of zero potential, so that one may then speak of the potential of a body. The surface of the earth, this being a conductor, is an equipotential surface [cf. p. 686] and is taken as the zero of reference. [Small variations of the potential of the ground due to earth currents, etc., are ignored.]

Analogies from other Branches of Physics.

(i) ***From hydrostatics.*** Suppose that A and B, Fig. 36·2, are two cans connected together by a pipe, fitted with a stop-cock, C,

as shown. Let this pipe be very small compared with A or B. Let equal quantities of water be placed in each can. The levels are indicated by the full lines. On opening C liquid passes from A to B until the levels are the same in each—shown by the dotted lines. The flow takes place in the direction from A to B since the pressure on the A-side of C is greater than that on the B-side.

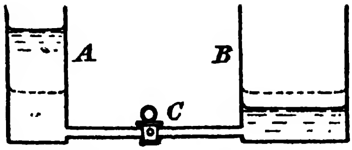


FIG. 36·2.—Potential-Analogy from Hydrostatics.

(ii) ***From Heat.*** Suppose a small iron ball is heated until it is red hot and then placed in contact with a large iron ball at room temperature. The large ball may contain more thermal energy or heat than the smaller one, the total energy associated with the molecules of a body being regarded as its heat content, but heat flows from the hot body to the cooler one, i.e. it is the temperature of a body which determines whether or not it shall communicate heat to another body in contact with it.

Electric Potential.—The potential of a body is its electrical condition determining whether or not electricity flows from it to earth, or vice versa, when there is metallic connexion between the body and earth. If the flow is from the body to earth, the potential of the body is said to be positive: it is negative when the flow is in the reverse direction.

The above is only a general statement about potential: to develop the theory of electrostatics it is necessary to have a precise definition of potential. Let us see how this is obtained.

Suppose that a small positive charge finds itself in an electric field—the presence of this charge will be assumed not to affect the original distribution of charges to which the field is due. Then owing to the existence of an electric intensity at the point where the charge is, it will experience a mechanical force. If the small test charge is not fixed it will move—in the direction of the electric intensity when the above charge is positive; in the reverse direction if it is negative. If the small test charge is moved about in the field, we may either have to do work against the forces due to the field, or they may do the work for us. Wherever the charge is situated there will be associated with it a definite amount of potential energy, this being equal to the work done in bringing the test charge from a point where the potential energy associated with it is zero to the point in question. This energy will be considered positive when the work is spent in overcoming the forces acting on the test charge. The potential energy will be zero when the electrical field strength is zero, i.e. at infinity.

If the electric field is due to a positively charged body, the electric intensity will be directed away from the body at all points in the field and work will be done in overcoming the force arising from the electric intensity at each point in the path of the test charge, when this is positive, as it moves from infinity towards the charged body. The potential energy of the test charge is everywhere positive. If the field is due to a negative charge, the potential energy is negative.

The above considerations enable us to define a difference of potential as follows.—***The potential at a point A exceeds that at a point B by the numerical value of the work done against the field in taking a unit positive charge from B to A.***

More strictly, ***the potential at a point A exceeds that at a point B by the work done against the field per unit positive charge in taking a small positive charge from B to A.***

In mathematical language we may state that the potential difference between A and B is the limiting value of $\frac{\delta w}{\delta q}$, i.e. $\frac{dw}{dq}$, where δw is the work done against the field in transferring a charge δq from B to A.

If B is at infinity the potential there is zero, so that ***the potential at a point in an electric field is equal to the work done against the field per unit positive charge in bringing a small positive charge from infinity to the point in question.***

It cannot be emphasized too strongly that whereas work, or potential energy, is measured in ergs (say), electric potential is measured in ergs per unit charge.

Definition of the c.g.s. electrostatic unit of potential difference. The potential at a point is one c.g.s. electrostatic unit of potential if one erg of work per unit charge is done against the field in bringing a small positive charge from infinity to that point.

This unit has no other name but it is equivalent to 300 volts, the volt being the practical unit for potential [cf. Chap. XLV].

The Earth's Gravitational Field.—The gravitational intensity or force per unit mass, near to the earth's surface is constant—it is equal to g absolute units of force per unit mass, since the force acting on a mass m is mg. Consequently, there must be a definite gravitational potential at points in the earth's gravitational field. It is because the gravitational intensity is constant at points near to the earth's surface that we are able to calculate the gravitational potential at points in that region. For if a mass m is raised through a vertical distance h the work done against the earth's field is mgh. Since this is the potential energy of the body of mass m at a height h, the potential at that point is gh.

From the above it will be realized that it is only because the

gravitational intensity is constant over the region considered that the gravitational potential at a point in that region can be calculated by very simple methods. In electrical fields the electric intensity is not, in general, constant, so that it is only in rather simple instances that the potential at a point in an electric field may be computed. In these calculations the zero of potential is selected to be at infinity: in practical problems involving the measurement of differences of potential, the earth is considered to be at zero potential.

The Principle of the Action of a Gold-Leaf Electroscope.—Suppose that a positive charge is given to an insulated metal body and this is allowed to touch the disc of a gold-leaf electroscope whose case is earthed, and therefore at zero potential, as in Fig. 36·3 (*a*). Electricity flows from the charged body to the disc and leaves, the charge on the leaves acting inductively on the case, where only a negative charge appears since the case is earthed. It will be found that the leaves diverge: there is a potential difference between them and the case.

If the charge on the body allowed to touch the electroscope is

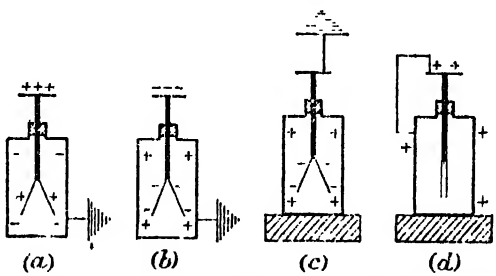


FIG. 36·3.—Principle of the Action of a Gold-Leaf Electroscope.

negative, the state of affairs is shown in Fig. 36·3 (*b*) when again the leaves diverge and there is a potential difference between them and the case.

Now let the electroscope be supported on an insulating stand and its disc earthed. When a positive charge is given to the case of the instrument a negative induced charge appears on the leaves which diverge: there is a difference of potential between the leaves and the case—cf. Fig. 36·3 (*c*).

When, however, the electroscope still being insulated, there is metallic connexion between the case and the disc, cf. Fig. 36·3 (*d*), no divergence of the leaves occurs when the case is charged: there is no difference of potential between the leaves and the case.

The above experiments show that the leaves of an electroscope will only diverge when there is a difference of potential between

them and the case of the instrument. If, as usual, the case is earthed, the divergence of the leaves measures the potential of the leaves and any body attached to them. [If the capacitance of the body (cf. p. 689) is large compared with that of the electroscope and charges are always communicated to the same body, then the deflexion of the leaves is a measure of the charge received by the body.]

Free and Induced Potentials.—Since work must be performed, either by the field or against it, when a charged body is brought from infinity to any point in the field of a charged body, it follows that a charged body must possess potential energy due to its own charge, for we may imagine that its charge has been obtained by bringing up in succession small charges from infinity, each process involving a certain amount of work on account of that fraction of the total charge already present on the body. The potential of a body due to its own charge is termed its *free potential.*

Now let it be supposed that an uncharged gold-leaf electroscope is brought into the field of a positively charged body for example. The electroscope finds itself in a field where the potential is everywhere positive: its leaves are deflected and therefore it must be at a definite positive potential itself. There is no charge on the electroscope as a whole, although from our study of electrostatic induction we know that negative electricity has been induced on the cap and positive on the leaves. The potential of the electroscope is termed an *induced potential.*

On Testing the Nature of a Charge by means of an Electroscope.—Suppose that an electroscope has been given a positive charge so that its free potential is positive. Let a positively charged body be brought gradually nearer to the electroscope. This will acquire, in addition to its own free positive potential an induced positive potential: since the instrument is a detector of potential differences the leaves will diverge further. The divergence increases as the distance between the body and the electroscope diminishes. This is a sure test for a positively charged body.

Now let a negatively charged body be brought near to a positively charged electroscope. This acquires a negative induced potential; its resultant potential is therefore reduced and the leaves collapse. As the body approaches more closely the divergence of the leaves decreases since the induced potential is becoming numerically greater. Ultimately, the leaves fall together and then begin to diverge again since the negative induced potential is numerically larger than the positive free potential of the instrument. The resultant potential is negative: this increases as the body is moved still nearer.

It must be borne in mind that an initial decrease in the diver-

gence of the leaves of a positively charged electroscope when a body is brought near, does not necessarily imply that the body is negatively charged for a similar effect is observed when the body is uncharged.

To test whether a body carries no charge at all an attempt is made to charge an electroscope by induction: if there is no induced charge the body under test is uncharged.

The Electrical Potential due to a Point Charge in Air.—Let q be a point charge situated at O, Fig. 36·4, and let P be the point at which the potential is required. Join OP and produce this line

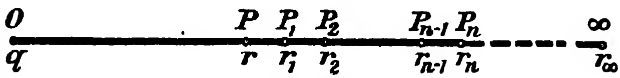


FIG. 36·4.

to infinity. Let P_1, P_2, P_3, . . . P_{n-1}, P_n . . . P , be points on this line at distances r_1, r_2, r_3, . . . r_{n-1}, r_n, . . . r_∞^∞ from O. Then V_P, the electric potential at P, is the work done against the field per unit charge in bringing up charge from infinity to P. Since the electric intensity, or force against which the work is done, is not constant between P and ∞, we have to proceed to calculate the work done as follows:—

The electric intensity at P is $\frac{q}{r^2}$; at P_1 it is $\frac{q}{r_1^2}$. Now if r and r_1 do not differ by more than a small amount, the arithmetical mean of these two quantities will be equal to their geometrical mean, viz. $\frac{q}{rr_1}$. Hence the work done against the field in bringing the small charge δq from P_1 to P will be

$$\frac{q}{rr_1}(r_1 - r).\ \delta q = q\left[\frac{1}{r} - \frac{1}{r_1}\right].\ \delta q.$$

Similarly the work done in moving the charge δq from P_2 to P_1 is $q\left[\frac{1}{r_1} - \frac{1}{r_2}\right].\ \delta q$. In general, the work done between two neighbouring points P_{n-1} and P_n is $q\left[\frac{1}{r_{n-1}} - \frac{1}{r_n}\right].\ \delta q$. Consequently the total work, W, done in bringing the charge δq from ∞ to P is

$$W = q\left[\left(\frac{1}{r} - \frac{1}{r_1}\right) + \left(\frac{1}{r_1} - \frac{1}{r_2}\right) + \ldots + \left(\frac{1}{r_{n-1}} - \frac{1}{r_n}\right) + \ldots\right].\ \delta q.$$

$$= q\left[\frac{1}{r} - \frac{1}{\infty}\right]\delta q = \frac{q}{r}.\ \delta q.$$

The work done against the field per unit charge is V_P. Thus

$$V_P = \frac{W}{\delta q} = \frac{q}{r}.$$

The above result may be obtained with the aid of the calculus as follows. At a point X, distance x from O where there is a point charge q, the electric intensity is $\frac{q}{x^2}$ and is directed along OX. Let OX be the x-axis and suppose a small positive charge is brought from infinity to X along this axis. If this charge moves from a point whose abscissa is $x + \delta x$ to X, the work done per unit charge against the field by the external agent is $-\frac{q}{x^2}\delta x$. The minus sign occurs since the distance moved is $-\delta x$. The work done per unit charge in bringing the charge from infinity to P, ($x = r$), is

$$-\int_{\infty}^{r} \frac{q}{x^2}dx = \left[\frac{q}{x}\right]_{\infty}^{r} = \frac{q}{r}.$$

If there are several point changes, $q_1, q_2 \ldots q_n$, the potential at a point P is $\sum_{n=1}^{n=n} \frac{q_n}{r_n}$ for the contribution from each charge to the total potential is unaffected by the presence of the other charges.

If P and Q are two points in an electrostatic field having potentials V_P and V_Q respectively [$V_P > V_Q$ (say)], then W, the work done by an external agent in moving unit-positive charge from Q to P, is numerically equal to $V_P - V_Q$. Since this value is independent of the potentials at points intermediate between P and Q it follows that the work done is independent of the actual path along which the unit charge is transported. If this were not so, energy could be obtained by taking a charge along one path and allowing it to return along another where less work was done when the charge was taken along it. This would violate the principle of the conservation of energy.

Also, since work = force $\times$ distance, the *average* electric field strength between P and Q, is $\frac{V_P - V_Q}{PQ}$, and is directed from P to Q if $V_P > V_Q$.

If P and Q are neighbouring points at potentials V and $V + \delta V$ respectively and at distances r and $r + \delta r$ from an origin O in QP produced, then, if E is the electric intensity at P in the direction of r increasing,

$-E \cdot \delta r$ = work done per unit positive charge by an external agent in taking a small positive charge from P to Q. The minus sign occurs since the work is done by the field. But the work done by the external agent is equal to the increase in potential in passing from P to Q, viz. δV.

$$\therefore \; -E \cdot \delta r = \delta V,$$

i.e. proceeding to the limit, $E = -\frac{\partial V}{\partial r}$.

The expression $\frac{\partial V}{\partial r}$ measures the ***potential gradient*** at P in the direction of r increasing.

Definition.—A surface in an electric field such that at every point on it the potential has the same value, is termed an equipotential surface.

Since no work is done when a charge is moved along an equipotential surface, it follows that the lines of force must be perpendicular to equipotential surfaces. To prove this, let E be the electric field strength at a point on an equipotential surface. Let V be the potential and let E make an angle θ with the tangent to the surface at the point considered. If δs is the small displacement of a unit positive charge from the above point to another on the same surface, the work done is

$$E \cos \theta . \delta s = 0$$

since $\delta V = 0$. If E is not zero, $\cos \theta = 0$, i.e. $\theta = \frac{\pi}{2}$.

Equipotentials due to Point Charges.—Since the potential at a point r cm. away from a point charge q is $\frac{q}{r}$, it follows that the equipotentials will be spheres having q at their common centre. A picture of these equipotentials is obtained by giving q some fixed value, and calculating the values of r corresponding to different potentials. Thus, for a point charge $q = 50$ e.s.u., the equipotential $V = 20$ is represented by a circle of radius $\frac{50}{20} = 2{\cdot}5$ cm. Fig. 36·5 indicates some of the equipotentials due to such a charge. It will be noticed that equipotentials differing by the same amount are nearer together the more closely they approach the charge. Since [average] electric intensity = potential difference ÷ distance, it follows that the intensity is greatest where such equipotentials are nearest together.

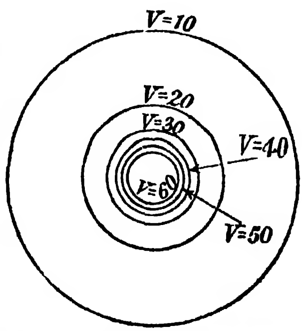


FIG. 36·5.—Equipotentials due to a Point Charge.

The equipotentials due to point charges q_1 and q_2 are obtained by drawing the series of equipotential surfaces due to each charge alone and drawing lines through those points which have the same resultant potential.

Exercise.—Construct the equipotentials and lines of force due to charges + 30 and − 10 e.s.u. at a distance apart of 5 cm.

The lines of force may be constructed by drawing curves cutting the equipotentials at right angles since the lines of force and equipotentials cut orthogonally, i.e. at right angles.

The lines of force and the equipotentials for an insulated sphere under the influence of a charge on a small body are shown in Fig. 36·6. It will be noticed that the lines of force meet the sphere at right angles, and that the equipotentials and lines of force intersect at right angles, i.e. they cut orthogonally.

[Although the lines of force and equipotentials have been drawn for a special instance, the nature of the curves is similar even when the shape of the conductor is more complicated.]

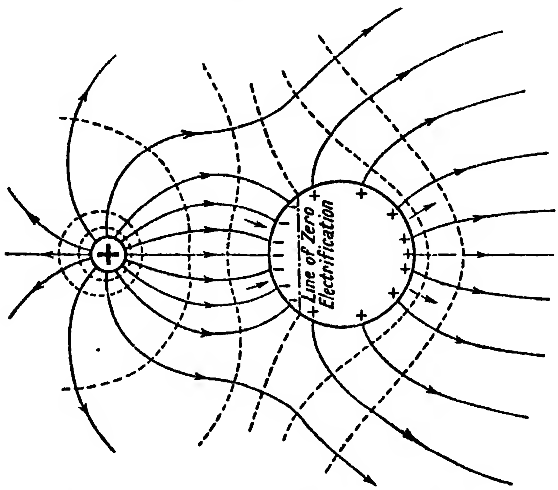


FIG. 36·6.—Lines of Force and Equipotentials for an Insulated Sphere under the Influence of a Small Charge.

Since the surface of the conductor is an equipotential with negative electricity on the side nearer the inducing charge and positive on the remote side, it follows that one of the equipotential surfaces must include the surface of the conductor itself. When only one kind of electricity is on a conductor, the equipotentials surround that conductor and one, of course, coincides with its surface.

The Potential due to a Uniform Distribution of Electricity on a Sphere.—Let P, Fig. 36·7 (*a*), be the point at which the potential due to a charge of surface density σ on a sphere of radius a is required. Let OP $= r$. If A and B are two points on the surface such that the angles AOP and BOP are θ and $\theta + \delta\theta$ respectively, the area of the ring traced out by AB when the figure is rotated about OP is $2\pi a \sin\theta \, a \, \delta\theta$. The charge on this is

$2\pi a^2\sigma \sin\theta\,\delta\theta$, and since each point on this ring is at the same distance x from P, the potential at P due to this charge is $2\pi a^2\sigma \sin\theta\,\delta\theta \div x$, where $x^2 = a^2 + r^2 - 2ar\cos\theta$.

Hence $x.\delta x = ar\sin\theta.\delta\theta$.

$$\therefore\ V = \frac{2\pi a\sigma}{r}\int_{x_1}^{x_2} dx,$$

where x_1 and x_2 are the values of x when θ is 0 and π respectively. For a point outside the sphere $x_1 = (r - a)$, $x_2 = (r + a)$.

$$\therefore\ V_e = \frac{2\pi a\sigma}{r}.2a = \frac{4\pi a^2\sigma}{r}$$

or if q is the total charge on the sphere, $V_e = \frac{q}{r}$.

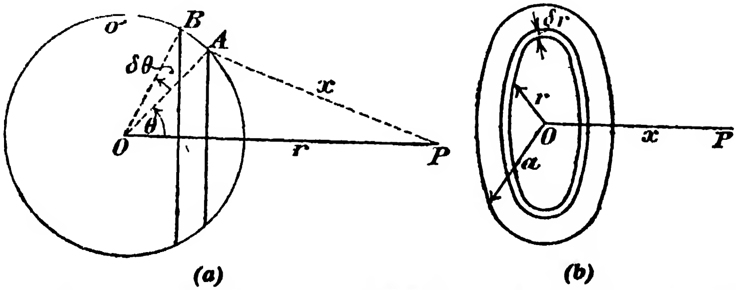


FIG. 36·7.—Potential due to (a) a Charged Spherical Conductor and (b) a Charged Metal Disc.

For a point inside the sphere $x_1 = a - r$, $x_2 = a + r$,

$$\therefore\ V_i = \frac{2\pi a\sigma}{r}.2r = 4\pi a\sigma = \frac{q}{a}.$$

i.e. the potential inside a charged sphere is everywhere constant and equal to that of the sphere itself. This last result is very important for it is true for all closed conductors [cf. p. 713].

Since $V_e = \frac{q}{r}$, the electric intensity outside is $-\frac{\partial V_e}{\partial r}$ or $\frac{q}{r^2}$, i.e., like the potential, it is the same as if the charge were concentrated at the centre of the sphere. Inside the sphere the intensity is zero, since the potential at points inside the sphere is constant and equal to that of the sphere itself. The result is true for all closed conductors.

The Potential at a Point on the Axis of a Disc having a Uniform Surface Charge of Density σ.—The charge on a ring whose inner and outer radii are r and $r + dr$, Fig. 36·7 (b), is $2\pi\sigma r.dr$. The potential at P due to this is $\frac{2\pi\sigma r.dr}{(r^2 + x^2)^{\frac{1}{2}}}$, since all points on the

ring are equidistant from P. If a is the radius of the disc, we have,

$$V = 2\pi\sigma \int_0^a \frac{rdr}{(r^2 + x^2)^{\frac{1}{2}}} = 2\pi\sigma . [(x^2 + a^2)^{\frac{1}{2}} - x].$$

Since the diagram is symmetrical about the x – axis, the resultant electric intensity must be directed along that axis. Hence

$$E = -\frac{\partial V}{\partial x} = -2\pi\sigma\left[\frac{x}{(x^2 + a^2)^{\frac{1}{2}}} - 1\right].$$

When the disc becomes very large, i.e. $a = \infty$, $E = 2\pi\sigma$, i.e. the electric intensity is constant.

In this problem we have assumed that the electricity is confined to one side of the disc. We know that it is on both sides so that the total electric intensity due to a uniform distribution on a metal disc is $4\pi\sigma$.

The Action of a Condenser.—Let A be an insulated metal plate which has acquired a positive charge and a corresponding positive potential. A second insulated plate B is then brought near to A so that B becomes charged by induction, the distribution of electricity being as in Fig. 36·8 (*a*). It is now necessary for us to consider the effect of B's charges on the potential of A. The negative charge on B will tend to diminish the potential of A, whilst the positive charge on B tends to increase it. Although the two induced charges on B are equal in amount the negative charge on B will have a greater effect on A than the positive charge will have, owing to the fact that it is nearer to A. Thus the sum total of the effect of B on A is to lower the potential of the latter. If A is connected to some constant source of potential, such as a battery of electric cells, then more electricity will flow from the battery to A in order to raise its potential to its original value V.

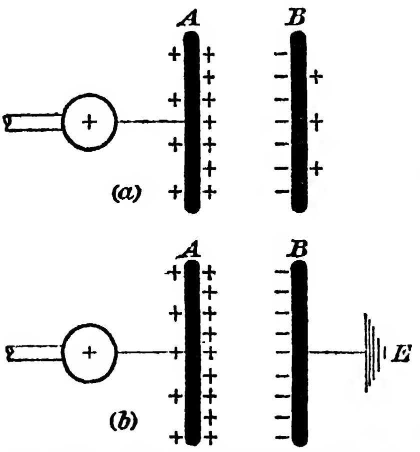


FIG. 36·8.—The Action of a Condenser.

When the plate B is connected to earth, Fig. 36·8 (*b*), there remains only the negative charge on it—in magnitude it is somewhat greater than before, since the nearby positive charge which

tended to diminish it has been removed.[1] The effect of this increased negative charge on A is to lower its potential again so that still more electricity flows from the battery to A. In other words the *capacity* of A for electricity has been increased.

Such an arrangement as this, in which an earthed plate is used to enable a second, but insulated, plate to acquire a greater charge, is called a plate condenser.

The *capacity,* or *capacitance,* (C) of a condenser is a constant for that condenser and is defined as ***the quantity of electricity on the positive plate per unit potential difference between its plates.*** Thus, if Q is the quantity of electricity on the positive plate and V the difference of potential across the condenser, then

$$C = \frac{Q}{V}.$$

Definition.—When one e.s.u. of charge raises the potential of a system by one e.s.u. of potential, the capacitance is said to be one e.s.u. of capacitance.

To Investigate the Factors influencing the Capacitance of a Condenser.—It will be assumed that the condenser consists of

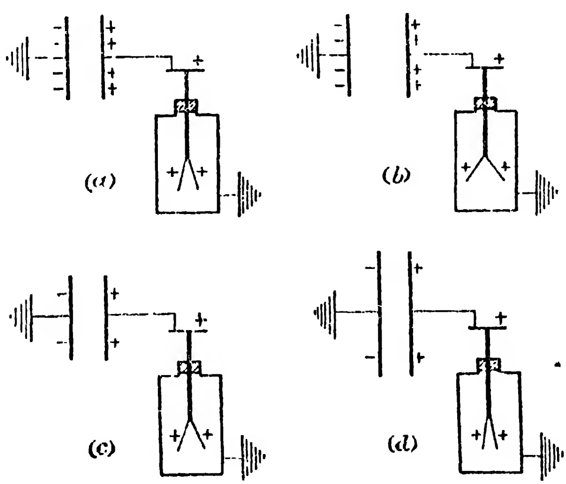


FIG. 36·9.—Factors Influencing the Capacitance of a Condenser.

two metal plates, one earthed and the other insulated and connected to the cap of an electroscope. Suppose that the insulated plate is charged positively—the portion of the charge on the electro-

[1] An analogy from the theory of magnetism may be useful here. It is difficult to obtain a strongly magnetized short needle because the two poles tend to neutralize one another.

scope is small and is not indicated quantitatively in the diagram. There will also be an induced charge on the earthed plate, but when we speak of the charge on a condenser we imply the charge on the insulated plate: numerically, the charges are, of course, equal. Such a system is shown in Fig. 36·9 (*a*). If the distance between the plates increases as in Fig. 36·9 (*b*), the angular separation of the leaves increases, i.e. the potential difference across the condenser has increased. Since the charge on the condenser has remained constant it follows that its capacity has diminished. If the plates are brought closer together the leaves tend to collapse, showing that the potential difference is less and the capacitance greater.

Suppose now that we have two condensers, the distance between the plates being constant but their areas different, the electroscopes to which they are connected being identical. Let equal charges

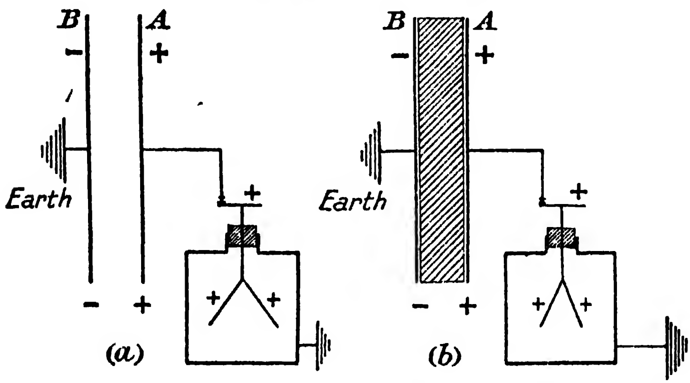


FIG. 36·10.—Action of a Dielectric (Insulator) on the Capacitance of a Condenser.

[cf. p. 677] be given to each system. The electroscopes will indicate that the potential of the (*c*) system is the greater, i.e. the capacity of (*d*) is greater than that of (*c*).

It now remains to examine how the capacitance of a condenser depends on the nature of the medium separating the plates. The insulating medium separating the plates of a condenser is referred to as a ***dielectric***—in the above the dielectric has been air. In Fig. 36·10 (*a*) the potential of a condenser is demonstrated by the divergence of the leaves of an electroscope. In 36·10 (*b*) a slab of insulating material, paraffin wax, ebonite, or glass, has been introduced between the plates. The leaves suffer a diminished divergence, showing that the potential of A has fallen. If A were connected to a source of constant potential then more electricity would flow to A, i.e. the introduction of the dielectric has increased the capacitance of the system.

The Electrostatic Capacitance of a Sphere in Air.—We have seen that the potential at a point in air (strictly in a vacuum) outside a charged sphere is equal to the charge on the sphere divided by the distance of the point from the centre of the sphere. The potential at a point on its surface due to the charge on the surface is therefore $\frac{q}{a}$, where a is the radius. There is also a charge $-q$ at infinity, but its contribution to the potential of the sphere is zero. The total potential is therefore $\frac{q}{a}$. The capacitance of the sphere is therefore $\frac{q}{V} = a$, i.e. the capacitance of a sphere expressed in c.g.s. electrostatic units of capacitance is numerically equal to its radius in centimetres.

Experiment.—A, Fig. 36·11, is a glass tube to which is attached a short metal tube B. C is a Woulf's bottle containing a small quantity of

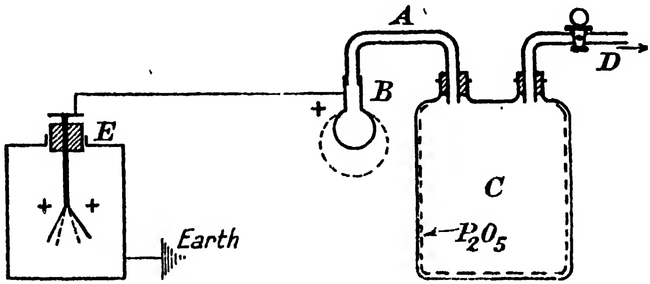


FIG. 36·11.—To show that the Capacitance of a Sphere increases with its Radius.

phosphorus pentoxide so that its interior shall be dry—to diminish the natural leak of the apparatus. D is a tube leading to a bicycle pump. B is joined to the disc of a gold-leaf electroscope E. A small soap bubble is blown at the end of B, and B is charged with electricity. The leaves of the electroscope diverge. When the bubble—shown dotted—is enlarged by forcing air into the apparatus the divergence of the leaves of the electroscope diminishes showing that the potential of the system has decreased—because its capacity has increased. If the air in C is allowed to escape, the bubble contracts and the divergence of the leaves increases.

The Capacitance of a Concentric Spherical Air Condenser.—Let A be a metal sphere and B a concentric spherical conducting shell connected to earth. Let a be the radius of A, and b the radius of the interior surface of B, for it is upon this surface that there will be an induced charge. Let Q be the charge on the inner sphere: suppose that Q_1 is the charge induced on the inner surface of the outer conductor. The potential of B due to the charge on

A is $\frac{Q}{b}$ since, at external points, the charge on a spherical conductor acts as if it were concentrated at its centre; its potential due to its own charge is $\frac{Q_1}{b}$, so that its total potential is $\frac{Q}{b} + \frac{Q_1}{b}$ and this is zero since the shell is earthed. Hence $Q_1 = -Q$. Now the potential of A due to its own charge is $\frac{Q}{a}$. But the potential at all points inside B, due to the induced charge on its interior, is $-\frac{Q}{b}$. Hence the actual potential of A is $Q\left[\frac{1}{a} - \frac{1}{b}\right]$. Since the outer sphere is earthed this expression gives the potential difference between the two spheres. The capacitance is therefore

$$\frac{Q}{V} = Q \div Q\left[\frac{1}{a} - \frac{1}{b}\right] = \frac{ab}{(b-a)}.$$

The Capacitance of a Parallel Plate Air Condenser.—When the radii of the above spheres are nearly equal, i.e. $b - a = t$, where t is small compared with a or b, the capacitance becomes $\frac{ab}{t} = \frac{a(a+t)}{t}$ $= \frac{a^2}{t}$ approximately. The capacitance per unit area of the condenser is therefore $\left(\frac{a^2}{t}\right) \div 4\pi a^2 = \frac{1}{4\pi t}$, a result independent of the radii of the spheres providing they are large compared with t.

When the radii of the spheres become very large and we consider unit area of the condenser, we may regard this as a condenser formed of two parallel plates each of unit area. Hence, if we have a parallel plate condenser formed by two very large plates each unit area of this condenser will have a capacitance equal to $\frac{1}{4\pi t}$ where t is the distance between the plates. If A is the area of each plate the capacitance of the whole will be $\frac{A}{4\pi t}$. This result is only true if t is small compared with the linear dimensions of the plates for it is only then that the lines of force are normal to each plate—in the spherical condenser considered above the lines of force are radial and therefore normal to the surface of each sphere. If t is increased the lines of force bulge outwards near the edges, an effect which increases with t, and the simple theory developed above is not applicable.

Condensers in Parallel and in Series.—To connect condensers in parallel their positive plates are joined together; likewise their negative plates. If we represent a condenser in the conventional

way by two equal straight lines, then Fig. 36·12 (*a*) shows three condensers in parallel. If c_1, c_2, and c_3 are the capacitances of the condensers, and q_1, q_2, and q_3 the charges on them when they are in parallel, while v is the common potential difference between the coatings of each condenser, then

$$V = \frac{q_1}{C_1} = \frac{q_2}{C_2} = \frac{q_3}{C_3} = \frac{q_1 + q_2 + q_3}{C_1 + C_2 + C_3}.$$

But since $q_1 + q_2 + q_3$ is the total charge on the compound condenser it follows that the capacitance of the latter is given by $C = C_1 + C_2 + C_3$.

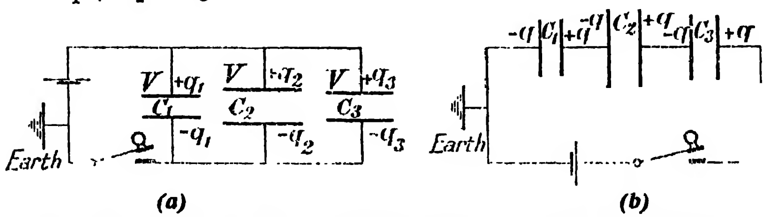


FIG. 36·12.—Condensers (*a*) in Parallel and (*b*) in Series (or Cascade).

By referring to Fig. 36·12 (*b*) we see how three condensers are arranged when they are in series. In this instance q is the numerical value of the charge on each plate. Let V_1, V_2, and V_3 be the potential differences between the plates of the condensers. Then

$$q = C_1V_1 = C_2V_2 = C_3V_3 = \frac{V_1 + V_2 + V_3}{\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}}.$$

But $V_1 + V_2 + V_3$ is the total potential difference, V, between the extreme plates of the compound condenser. Hence

$$q = C(V_1 + V_2 + V_3),$$

where C is its capacitance, so that $\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}$.

Dielectric Constant.—When an insulator (or dielectric) is inserted between the plates of a condenser it has been shown that the capacitance is increased. If the dielectric completely fills the space between the plates of the condenser, then it is found that the ratio

$$\frac{\text{Capacitance of condenser with a dielectric}}{\text{Capacitance of same condenser in air}} = \text{constant } (\kappa).$$

This constant is known as ***the specific inductive capacity,*** the ***dielectric constant,*** or the ***permittivity*** of the insulator with reference to air, whose dielectric constant is unity for all practical purposes.

The Leyden Jar.—The Leyden jar, so named after the city where it was invented, is a common form of condenser. In its usual form it consists of a glass jar lined inside and out to perhaps three-quarters of its height with tinfoil. A brass knob, Fig. 36·13 (*a*), is attached to one end of a brass rod passing through a wooden lid fitting the mouth of the jar. This wood is well covered with shellac varnish to improve its insulating properties. Electrical communication between the knob and the inner coating of the jar is made by a short length of brass chain hanging from the lower end of the rod.

The outer coating of the jar is earthed. Hence, when the knob is connected to a source of positive electricity a positive charge is acquired by it and the inner coating of the jar. This charge acts inductively on the outer coating, and since this is connected to earth, a negative charge is retained by it. Such a jar may be regarded as a plate condenser with glass as the dielectric.

Let h be the height of the tinfoil, r the radius of the base, and t the thickness of the glass. Then the condenser consists of two condensers in parallel [cf. p. 692]—

(i) the cylindrical condenser of area $2\pi rh$, thickness t,

(ii) a parallel plate condenser of area πr^2, thickness t, the dielectric constant of the medium between the plates being κ in each instance.

The total capacitance of the jar is therefore

$$\kappa\left[\frac{2\pi rh}{4\pi t}+\frac{\pi r^2}{4\pi t}\right]=\kappa\left[\frac{rh}{2t}+\frac{r^2}{4t}\right],$$

if end effects are neglected.

The advantage of a Leyden jar is not that its capacitance is large, it is not, but that it will withstand large potential differences without breaking down, i.e. its dielectric is not punctured easily. Moreover, its capacitance does not fluctuate. In order to render the dielectric less susceptible to the formation of a puncture when there is a high potential difference across it, the glass must be of a good quality—air bubbles in the glass must be avoided since they diminish the mean dielectric strength of the glass (see later). Moreover, the outside of the jar should be coated with shellac varnish to reduce the leakage of electricity over the surface by retarding the deposition of moisture and permitting dust to be removed easily. Both dust and moisture increase the rate at which electricity leaks across a surface.

Glass is not punctured easily by an electric spark passing through it when placed in a strong electric field, because it possesses great ***dielectric strength,*** the latter being defined as the volts per cm. of thickness necessary to cause a breakdown in the material by sparking through it. [1 volt = $\frac{1}{300}$ e.s.u. of potential difference.]

TABLE OF DIELECTRIC STRENGTHS

Material.	Dielectric strength in volt.cm.$^{-1}$.
Ebonite	500,000
Glass	300,000
Mica	600,000
Paraffin oil	80,000
Air (S.T.P.)	40,000

Since the dielectric strength of a material is not a constant, but decreases with increasing thickness, the above table is not complete without the statement that the thickness of the substance when the above tests were made was 1 mm.

[Strictly speaking, the above definition of dielectric strength is not exact, since it is found that the facility with which a break-down occurs depends on the curvature of the terminals between which the electric field is produced.]

From the above we see that although the capacity of a Leyden jar may be increased by decreasing the thickness of the glass, in practice, this thickness is seldom less than 2 mm. owing to the fact that thin layers are more readily penetrated by an electric spark, and, of course, a thin-walled jar is more easily broken by accidental mechanical shocks.

Other Types of Condenser.—For condensers of the order of magnitude 1 μF, i.e. 1 microfarad [cf. Chap. XLVI], mica is the best material to be used as the dielectric. It may be split into very thin sheets since the material has a natural cleavage plane, and it has good mechanical properties. The sheets are assembled with pieces of tinfoil interposed between adjacent mica sheets. The alternate sets of tinfoil are soldered to copper leads connected to well-insulated terminals, and the whole covered with hot melted paraffin wax free from moisture. As soon as possible the condenser is sealed in an air-tight case: otherwise the paraffin absorbs water and the insulation resistance of the condenser is impaired.

If there are (N + 1) metal sheets in a condensor of this type, then the number of slabs of dielectric in a state of electrical strain is N: thus the capacitance of the condenser will be N times that of a 'single component', i.e.

$$C = \frac{\kappa AN}{4\pi t} \text{ e.s.u.,}$$

where the symbols have their usual significance.

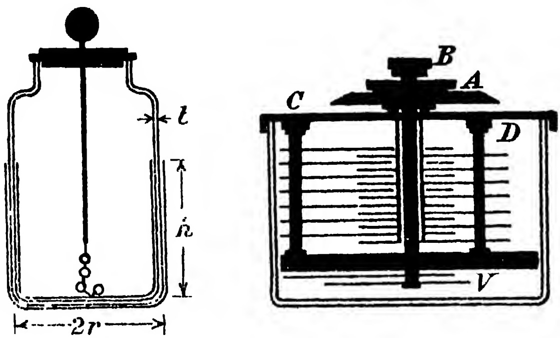
Example.—If N = 21, A = (18 × 10) cm.^{2}, t = 0·2 mm. and κ = 6·28, then

$$C = \frac{6{\cdot}28 \times 180 \times 20}{4 \times 3{\cdot}14 \times 0{\cdot}02} = 9 \times 10^4 \text{ e.s.u.}$$

In the practical unit of capacitance [cf. Chap. XVLI], this is

$$\frac{9 \times 10^4}{9 \times 10^5}\ \mu\mathrm{F} = 0{\cdot}1\ \mu\mathrm{F}.$$

The variable condensers which are employed for wireless purposes consist of two sets of brass or aluminium plates. The fixed set are semicircular in shape; the moving plates are shaped like a cam [i.e. half a heart] and operated by the knob A. The fixed plates



(a) A Leyden Jar.

(b) A Continuously Variable Condenser with Vernier Adjustment.

FIG. 36·13.

are all connected to one terminal, C, of the condenser, and the set of moving plates to the other terminal, D, air being the dielectric [cf. Fig. 36·13 (*b*)]. V is a 'vernier,' i.e. a small condenser in parallel with the condenser and operated by turning the knob B.

Boxes of Standard Condensers.—Various devices are used in practice for connecting condensers, either in parallel or in cascade (series), when the condensers are mounted in box-form. In the first arrangement, Fig. 36·14 (*a*), A, B, C, D, E, and F are brass bars outside the box and connected to the condensers as shown. X and Y are two brass bars placed at right-angles to the others. Plugs inserted in tapering holes $a - f$ and $a_1 - f_1$ enable X or Y to be put into connexion with any of the other bars. Such an arrangement enables the component condensers to be connected in parallel or in cascade. T_1 and T_2 are terminals.

Let us suppose that plugs, represented by the black dots, have been inserted as shown. Then the condensers between A and B, and E and F are out of action, while the condenser between B and C, is in parallel with a condenser consisting of the condensers between

C and D, and between D and E, arranged in cascade. The combined condenser therefore has a capacitance

$$0{\cdot}1 + \frac{1}{\left(\frac{1}{0{\cdot}2} + \frac{1}{0{\cdot}5}\right)} = 0{\cdot}243\mu F.$$

The key K_1 enables the condensers to be charged from the battery, K_2 being open, and when K_1 is open and K_2 closed the condensers are discharged through the ballistic galavanometer G.

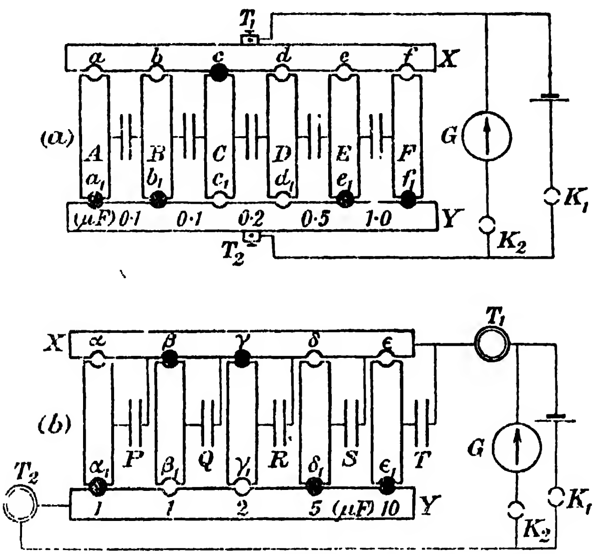


FIG. 36·14.—Boxes of Standard Condensers.

If the six plugs belonging to this box are arranged so that A is connected to X, B to Y, C to X, etc., they are said to be staggered and the combined capacity is 1·9 μF.

Precautions should be taken never to arrange two plugs so that any bar, such as D for example, is connected both to X and Y at the same instant when K_1 is closed.

Another method of mounting condensers in box-form is indicated in Fig. 36·14 (*b*). In this method the individual condensers may only be arranged in parallel. The internal connexions are shown in the diagram. When a plug is inserted in the hole γ, the condenser R is short circuited since both sets of plates are connected to X: when plugs are placed in the holes α_1, β, γ, δ_1, and ε_1, the condensers P, S, and T are in parallel, the total capacitance being $P + S + T = 16\ \mu$F, if the individual condensers have the capaci-

tances indicated. When this box is in use, all the plugs should be inserted in one or other of the holes so that even the condensers which are not in the circuit shall be definitely short circuited. This prevents any indefinite inductive action between the charges on the condensers in use and the other condensers. T_1 and T_2 are the terminals, and the condenser is charged and discharged as in the previous arrangement.

A Guard-Ring Condenser.—A precision type of guard-ring condenser of the parallel plate type is shown in Fig. 36·15. AB is one plate of the condenser, DE the other, and this is surrounded by the guard-ring GG. The ring forms part of a cylindrical box and the plate DE is insulated from the rest of the box by the quartz insulators QQ. The annular space between DE and G is very small. AB is earthed, while DE and G are connected to

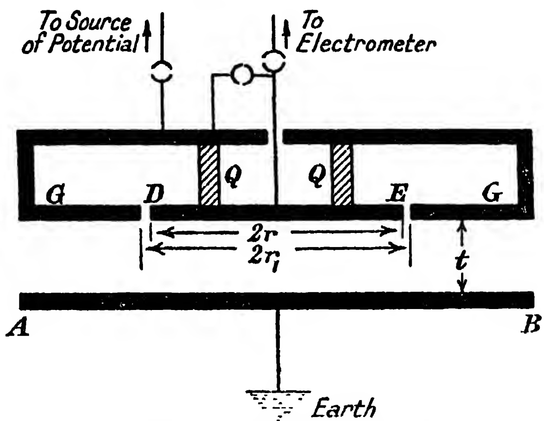


FIG. 36·15.—A Guard-ring Condenser.

the same source of potential, the connexions to each then being broken. If $2r$ is the diameter of DE and $2r_1$ that of the aperture in G, t the distance between the plates, the capacity, C, of DE and the portion of the plate AB directly in front of it is given by the relation

$$C = \frac{1}{4\pi t}\text{(mean area of plate DE and the aperture in G)}$$

$$= \frac{1}{4\pi t}\frac{\pi(r^2+r_1^2)}{2} = \frac{r^2+r_1^2}{8t}.$$

The advantages of this type of guard-ring condenser are—(i) that the whole of the charge on DE resides on its outer face, and its inner face is screened from outside charges, (ii) the portion of the field between DE and the other plate of the condenser is uniform.

The Potential Energy of a Charged Condenser.—Let us suppose that a condenser has a charge Q and that V is the potential difference between its plates. The potential energy of such a charged condenser is equal to the work done in bringing up the charge Q from infinity so that the final potential is V. This must not be confused with the potential at a point, which is numerically equal to the work done in bringing up unit-positive charge to the point in question. Suppose that the condenser is charged by bringing up small charges in succession. Then since the potential of a conductor is proportional to the charge on it provided it is at a great distance from all other bodies, the co-ordinates of any point on the straight line OV, Fig. 36·16, represent the charge and corresponding potential of this conductor at a particular instant. Let L be such a point. In bringing up a small charge so that the charge is increased from OA to OB, the work done is LA . AB which, when AB is small, is equal to the area of the trapezium LMBA. Similarly the area of the trapezium MNCB gives the work done in bringing up the next small charge. The total work done in bringing up the charge Q is therefore given by the area of the ΔOQV. Hence the potential energy, W, of the charged conductor is $\frac{1}{2}$QV. Since Q = CV this equation may be written $W = \frac{1}{2}\frac{Q^2}{C}$, or $W = \frac{1}{2}CV^2$. [Cf. Chap. XLVI, for note on units.]

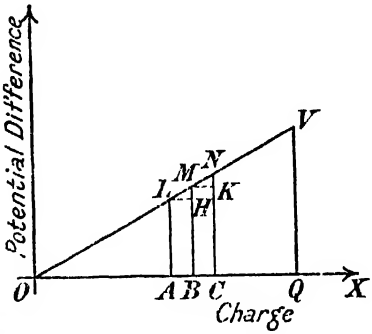


Fig. 36·16.

Alternative Proof.—Let us suppose that at some instant there is a charge q on the conductor and that its potential is v. If a charge δq is then added, the work done against the field is $v.\delta q$, for the potential may be regarded as constant. Hence, W, the total work done against the field in bringing up the charge Q, is given by

$$W = \int_0^Q v.dq = \int_0^Q \frac{q}{C}.dq = \frac{1}{2}\frac{Q^2}{C}.$$

If $V = \frac{Q}{C}$ is the final value of the potential of the charged conductor, we have

$$W = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV = \frac{1}{2}CV^2.$$

This work measures the potential energy of the charge on the conductor and appears as some other form of energy when the condenser is earthed.

The Loss of Energy on Connecting Two Charged Condensers in Parallel.—If two condensers of capacity C_1 and C_2 respectively and having charges Q_1 and Q_2 are connected in parallel there is no loss of charge and the potential difference between the plates of each condenser has the same final value. The following analysis shows, however, that there is a loss of potential energy. This appears as the energy of the spark in the form of light and sound energy, which is finally converted into heat energy. Before connecting the condensers in parallel the total energy is

$$\frac{1}{2}\frac{Q_1^2}{C_1} + \frac{\frac{1}{2}Q_2^2}{C_2}.$$

After connecting them it is $\frac{1}{2}\frac{(Q_1 + Q_2)^2}{C_1 + C_2}$. Hence the loss in energy is

$$\frac{1}{2}\left[\frac{Q_1^2}{C_1} + \frac{Q_2^2}{C_2} - \frac{(Q_1 + Q_2)^2}{C_1 + C_2}\right] = \frac{1}{2}\frac{(Q_1C_2 - Q_2C_1)^2}{C_1C_2(C_1 + C_2)}.$$

There will therefore be no loss in potential energy if $Q_1C_2 - Q_2C_1 = 0$ or $\frac{Q_1}{C_1} = \frac{Q_2}{C_2}$, i.e. if the two condensers are at the same potential before they are connected in parallel.

Example.—A Leyden jar has a diameter 10·4 cm., while the glass is 0·25 cm. thick. If the height of the cylindrical coating is 20·5 cm. and the dielectric constant of glass is 6 calculate its capacity.

This jar may be regarded as a plate condenser the area A of which is equal to the sum of the areas of the base and sides. Hence

$$A = [\pi \times (5{\cdot}2)^2] + [2 \times \pi \times 5{\cdot}2 \times 20{\cdot}5]\ \text{cm.}^2 = \pi \times 240{\cdot}2\ \text{cm.}^2.$$

Hence

$C = \frac{\kappa A}{4\pi t} = \frac{6 \times 240{\cdot}2}{4 \times 0{\cdot}25} = 1{,}440$ e.s.u. $= 1{\cdot}6 \times 10^{-3}$ microfarad [cf. Chap. XLVI].

Example.—Two condensers of capacitance 30 and 23 e.s.u. respectively have charges + 8 and − 6 e.s.u. Determine the loss in potential energy when they are connected in parallel.

Total energy before connecting $= \frac{1}{2}\cdot\frac{64}{30} + \frac{1}{2}\cdot\frac{36}{23} = 1{\cdot}85$ ergs. After connecting the total charge is + 2, while the total capacity is 53 cm. The energy is then $\frac{1}{2}\cdot\frac{4}{53} = 0{\cdot}04$ erg. The loss in potential energy is therefore 1·81 erg.

The Seat of Electrical Energy.—FRANKLIN discovered by means of the following experiment that when, for example, a Leyden jar is charged the electrical energy is stored in the glass. The type of jar he used had detachable coatings as illustrated in Fig. 36·17. The jar is assembled and then charged. If the inner coating is removed with the aid of insulated tongs and the glass carefully lifted out no charge will be found on either of the coatings. Yet when the parts are reassembled a vigorous spark is obtained when the inner and outer coatings are joined together. This indi-

cates that the energy is stored in the dielectric [glass] between the metal coatings. The glass is said to be ***strained*** (electrically). If the glass wall of a Leyden jar is thin and the charge high the strain in the glass may become so great that the glass fractures.

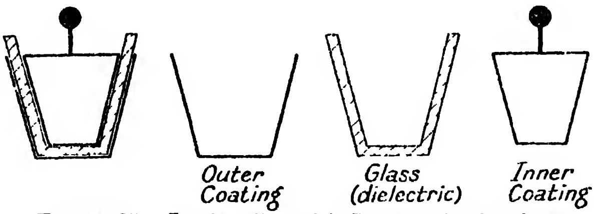


Fig. 36·17.—Leyden Jar with Detachable Coatings.

The following experiment shows that the electricity does not reside on the surface of the glass: while the jar is dismantled X-rays are allowed to fall on the glass—X-rays are able to remove surface charges. Yet when the jar is re-assembled a spark is obtained when the inner and outer coatings are joined together.

The Residual Charge.—The fact that the dielectric of a condenser is strained accounts for the following phenomenon. If a Leyden jar is allowed to rest after being discharged and its coatings then brought into conducting communication by means of discharging tongs, a spark will often pass. This is because the glass does not recover itself at once after being strained.

The Action of Points.—In an earlier chapter, cf. p. 675, it was shown that the surface density of electricity on a conductor is greatest where the radius of curvature of the surface is least. When the radius becomes very small there is a 'point' on the conducting surface and it is here that the surface density of the electricity is greatest and may reach a high value. Experimentally it is found that an insulated conductor with a point on its surface rapidly loses its charge, a fact first noted by FRANKLIN. The reason for the above loss is that those air particles which come into contact with the point acquire a similar charge—it is taken from the point—and are then repelled. Of course this electrification by 'direct bombardment' takes place at all portions of the surface, but more air particles are likely to acquire a charge at those portions of the surface where the density of the charge is greatest. If the conductor is connected to a Wimshurst machine [cf. p. 736] the air molecules are repelled so violently that a wind is produced. This wind may be sufficient to extinguish a lighted candle.

FRANKLIN also demonstrated the action of points by connecting a small 'wind-mill' to an electrostatic machine. This mill consisted of several wires attached radially to a metal axle, the whole

being placed on a pivot. The ends of the wires were all bent to point in a clockwise-direction, say; it was found that the mill *rotated in an anticlockwise direction when it was charged continuously.* This apparatus is a simple reaction turbine whose motion is produced by the repulsion of charged molecules of the constituents of air away from the several points.

Points also enable electric charges to be removed from charged conductors without any actual contact between the point and conductor taking place. Suppose A, Fig. 36·18 (*a*) is a charged conductor. When an uncharged conductor B, with a point on it, is placed as indicated near to A induced charges appear on B and are distributed as shown in Fig. 36·18 (*b*). The negative charges at the point are acquired by the air molecules as they come into contact with this part of the conductor and then, being negatively

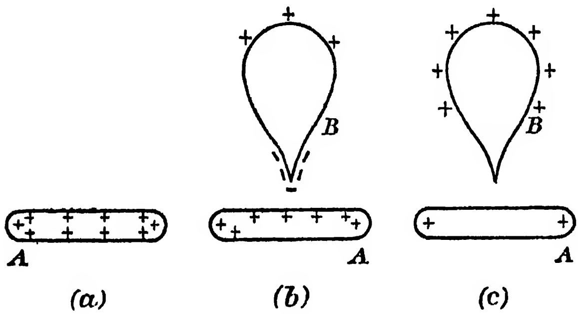


FIG. 36·18.

charged bodies, are attracted to A so that part of the charge on A is neutralized. This continues until A and B are at the same potential when it is found that A has lost its charge almost completely and there is a large excess positive charge on A—cf. Fig. 36·18 (*c*).

In addition it must be remembered that in the vicinity of a point the strength of the electric field is high so that the gas becomes ionized, i.e. there are created in the gas positively and negatively charged systems known as gaseous ions. These consist of atoms which have gained or lost an electron. The positively charged ions are attracted to the point and neutralize some of the negative charge on it: the negative ions move towards A so that it, too, loses some of its charge. Thus at least two processes may be at work whereby the charge is transferred from the body A to another conductor with a prominent point or 'spike' on it.

The action of a lightning conductor is due to the discharging action of the points at its top. When a cloud, generally with a positive charge on its lower surface, appears over any particular

region of the earth's surface a negative charge will be induced on the earth. A large potential difference will be established between earth and cloud. When this p.d. is sufficiently large the insulation of the air breaks down and a discharge—the so-called lightning flash—occurs. If, however, there is present in the region below the charged cloud a building provided with a lightning conductor, the density of the negative electricity will be greatest at the tip of the conductor. This negative charge is removed by the air molecules which, when charged, move towards the cloud and neutralize a portion of the charge on it. In this way the cloud is discharged continuously and the potential difference between cloud and earth is kept low so that no visible discharge occurs.

EXAMPLES XXXVI

1.—A spherical conductor has a charge of 843 e.s.u. when its potential is 250. What is the radius of the sphere? Assuming the charge to be distributed uniformly over the surface, calculate the charge per unit area.

2.—Two spheres of radii 3 and 5 cm. respectively are each given a charge of 30 positive units. If the spheres are then connected by a wire, calculate their common potential in e.s.u. and in volts.

3.—A condenser of capacitance 84 units has a potential of 2000 e.s.u. When its charge is shared with a spherical conductor the potential is 1500 e.s.u. What is the radius of the sphere?

4.—A soap bubble has a charge of 64 e.s.u., its radius being 8·5 cm. What is the change in potential when the radius is increased by 1 cm.?

5.—Define *electric potential* and explain how it is measured. How would you show experimentally that it is possible for parts of the same conductor to be oppositely electrified and yet at the same potential?

6.—What is implied by the statement '*the dielectric constant* (specific inductive capacity) of glass is 6 and that of ebonite is 2'? Describe how you would compare the dielectric constants of glass and ebonite if these substances were available in the form of sheets each 2 cm. thick.

7.—Explain why a 'Leyden jar' is described as a condenser. Two insulated metal plates 20 cm. in diameter, having opposite charges of 5 electrostatic units each, face each other across a layer of air 2 mm. thick. Calculate the potential difference between them and the electrical energy of the system.

8.—Describe how you would investigate whether the electrical capacitance of an insulated conductor depends upon other bodies in its neighbourhood.

9.—Calculate the capacitance of a parallel plate condenser, the medium between the plates being uniform and having a dielectric constant κ. If such a condenser has a charge 50 e.s.u. calculate the energy dissipated when its plates are connected together, if each plate has an area 20 cm.^{2} and that the interval between them is 1 mm. Assume κ to be 2·5.

10.—What is meant by electrostatic potential? Define the unit in which it is measured. Show that the capacitance of a parallel plate

condenser is given approximately by the expression $\frac{\kappa A}{4\pi d}$, where A is the area of each plate, d their distance apart, and κ the specific inductive capacity of the medium between the plates. Why is the expression approximate only ?

11.—Define the electrostatic units of quantity, potential difference, and capacitance. Obtain expressions for the capacitance of (a) an isolated sphere, (b) a parallel plate condenser.

12.—An insulated metal sphere is placed between, and near to, two similar spheres, one of which is positively charged and the other earthed. Draw a diagram illustrating the distribution of the lines of force, and discuss the potential of the different parts of the system.

13.—Define *electric potential* and explain what is meant by an *equipotential surface.* A charged sphere is placed near to an insulated uncharged sphere of the same size. Draw a diagram illustrating the positions of equipotential surfaces.

14.—What is meant by the electrical capacitance of a system ? How would you investigate the effect on the capacitance of a parallel plate air condenser of (a) increasing the distance apart of the plates, (b) filling the space between the plates with paraffin wax instead of air ? What results would you expect to obtain ?

15.—Describe a Leyden jar and derive an approximate expression for its capacitance. If you were provided with two such jars, a source of constant potential, and a gold leaf electroscope, how would you determine which jar had the greater capacitance ?

16.—A condenser of capacitance 10 microfarads is charged so that the potential difference between its terminals is 50 volts. The terminals of an uncharged condenser of capacity 2·5 microfarads are then connected to those of the charged condenser. Calculate : (a) the energy in the larger condenser before it is connected to the smaller ; (b) the potential difference between the terminals of the combination ; (c) the sum of the energies in the two condensers after they are connected to one another. (N.H.S.C. 29.)

17.—Find expressions for the electrical capacity of a sphere of radius r, and of a condenser the plates of which are concentric spheres of radii r and $(r - t)$. Hence find the capacitance per unit area of a plate condenser of thickness t, the dielectric being air in each case. (L. '24.)

18.—A spherical air condenser consists of an insulated metal sphere of radius 3 cm., surrounded by a concentric hollow metal sphere of internal radius 4 cm. and external radius 5 cm. If a charge of 100 e.s.u. is given to the inner sphere calculate the potentials of each sphere when the outer one is (a) insulated, (b) connected to earth. (L. '30.)

CHAPTER XXXVII

THE THEORY OF ISOTROPIC DIELECTRICS

The Experimental Basis for the Theory.—Let us suppose that we have two condensers geometrically alike, but that one is filled with an insulating material such as sulphur. If both are charged by being connected to the same high potential battery, as suggested in Fig. 37·1, and then discharged in turn through a ballistic galvanometer, the throws of the galvanometer are not equal, but their ratio is constant however the potential difference common to each condenser is varied. The ratio of these throws is numerically equal to the ***dielectric constant*** or ***dielectric coefficient*** or ***permittivity*** of the insulator. It is also equal to the ratio of the capacity of the condenser with the dielectric to that of the condenser with air (strictly a vacuum). [We assume that the condensers are long and narrow in order to render the end effects negligible.]

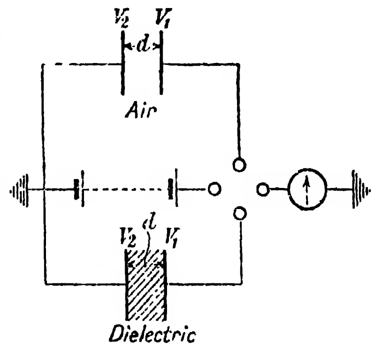


FIG. 37·1.—Geometrically Identical Condensers with and without a Dielectric.

If d is the distance apart of the plates in each instance, V_1 and V_2 the potentials of the plates of the condensers, then

$$\frac{V_1 - V_2}{d}$$

is the same for each. For the air condenser this expression measures the electric intensity in the electric field between the plates. What does it measure when the space between the plates is filled with an insulator? Before this can be answered the following digression is necessary.

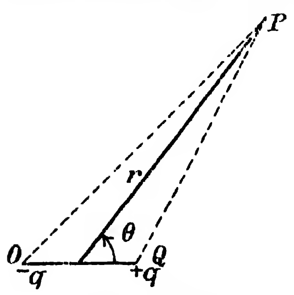


FIG. 37·2.—Potential at a Point in Air due to an Electric Dipole.

Electric Doublets or Electric Dipoles.—If two electric charges, equal in magnitude but opposite in sign, coincide at any point, the electric intensity due to these charges in the space round them must vanish. If, however, the charges suffer a small relative displacement any other charge introduced in the region round

them experiences a mechanical force, i.e. there is an electric field of sensible magnitude in the neighbourhood of the two charges. Such a combination of electric charges is termed an ***electric doublet*** or ***dipole***.

We shall see later that the molecules of substances such as ammonia, NH_3, water, H_2O. and chloroform, $CHCl_3$. are permanent dipoles; other molecules such as those of methane, CH_4, argon, A, oxygen, O_2, etc. are not permanent dipoles: they are said to be non-polar.

The Electric Field in Free Space due to a Dipole.—Let OQ, Fig. 37·2, be an electric dipole. Let OQ $= 2l$. Consider the potential, V_P, at a point P in free space distance r_1 from Q and r_2 from O. Then,

$$V_P = \frac{q}{r_1} - \frac{q}{r_2} = q\left[\frac{1}{r - l\cos\theta} - \frac{1}{r + l\cos\theta}\right], \text{ [if OQ is small.]}$$

$$= \frac{q.2l\cos\theta}{r^2} = \frac{q.\text{OQ}\cos\theta}{r^2}, \text{ [if } l \text{ is small compared with } r.]$$

$$= \frac{\mu\cos\theta}{r^2},$$

if $\mu = q.\text{OQ}$, the so-called ***electric moment of the dipole***.

The electric intensity in the direction of r increasing is

$$-\frac{\partial V_P}{\partial r} = \frac{2\mu\cos\theta}{r^3}.$$

The field at right angles to this and in the direction of θ increasing is

$$-\frac{1}{r}\frac{\partial V_P}{\partial\theta} = \frac{\mu\sin\theta}{r^3}.$$

Electric Polarization.—According to modern electrical theory an atom in its normal state is a configuration of electrons (negative charges)

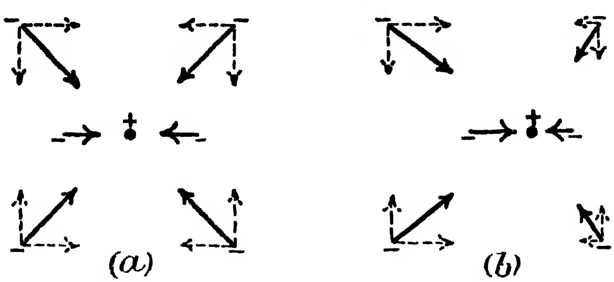


Fig. 37·3.—Electric Moments of Atoms.

surrounding a relatively massive nucleus carrying a positive charge numerically equal to the total charge of the electrons. Hence, as a whole, the atom is uncharged. For the present we shall adopt the 'static' atom, i.e. we shall consider the electrons to be stationary relative to the nucleus. Since the total charge of a normal atom is zero, we may pair off each electron with an equal positive charge on the nucleus and regard each normal atom as an assemblage of electric doublets. On the whole the total electric moment is zero. It is known that a carbon atom has six extra-nuclear electrons and a corresponding positive charge on the nucleus. If we imagine the electrons to be coplanar, then the electric doublets formed in the manner described above are as indicated in Fig. 37·3 (*a*). When the above configuration

of charges is subjected to an electric field, the negative charges are displaced, relative to the nucleus, as in Fig. 37·3 (*b*). Now an electric moment is a vector represented by the straight line joining the charges. Hence, like any other vector quantity, it may be resolved into components. The resolved components in the present instance in the direction of the field and perpendicular to it are indicated by the dotted lines. These show that as a whole the atom has acquired an electric moment in the direction of the field. When the above occurs with each of the constituent atoms of a dielectric, the medium of the dielectric is said to be ***polarized.***

If we take a microscopic view of a dielectric and imagine a being sufficiently small to wander in and out among the atoms and to be provided with suitable measuring instruments, then such a being would be able to detect variations in the electric field in the space between the different electric charges. Actual instruments are only able to detect the combined effect of an exceedingly large number of atoms and their charges, and cannot reveal the changes which do occur locally in the medium. Such instruments only indicate mean values. In an unpolarized medium, if we assume a chaotic arrangement of atoms, the resultant electric moment per unit volume must be zero. When a dielectric is subjected to an electric field there is a tendency for each atom to acquire a resultant electric moment in the direction of the field. i.e. there will be a finite electric moment per unit volume. This is termed the ***polarization*** of the medium and is denoted by the letter P.

So far, reference has only been made to atoms or molecules which become polarized when acted upon by an external electric field. There are substances, however, whose molecules are permanently polarized —the case is analogous to that of paramagnetism—and the effect of an electric field is to cause the molecules of such substances to align themselves with their axes parallel to the direction of the field, in addition to a change in the electric moment of each atom due to the displacement of its electrons relative to its nucleus. Later on, we shall see how experiment is able to discover the type of molecule present.

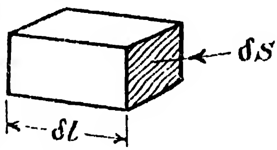


FIG. 37·4.—Electric Polarization.

Now let us consider a small element of a polarized medium; let δl be the length and δs the cross-sectional area of this element, see Fig. 37·4. Although this element is small it must still be sufficiently large to contain a large number of atoms so that the resultant electric moment of all the atoms in this element is parallel to the length of the element which is supposed to be in the direction of the applied electric field. The electric moment of this element is by definition $P.\delta s.\delta l$. Now this moment may be considered to arise from charges numerically equal to δq on the ends of the element. Its moment is then $\delta q.\delta l$. Equating these two expressions for the electric moment of the element, we have

$$P.\delta s = \delta q$$

or

$$P = \lim \frac{\delta q}{\delta s} = \frac{dq}{ds}.$$

Hence P, the polarization of the medium, is equal to the surface density of the electrification arising on the ends of an element such as that we have considered.

Electric Displacement or Electric Induction.—Let us consider unit cross-sectional area of a parallel plate condenser, (i) when the medium is air, (ii) when the medium has a dielectric constant κ. If V is the difference in potential between the plates, and σ_1 the surface density of the electrification, then

$$4\pi\sigma_1 = E = V/d \qquad . \quad . \quad . \quad . \text{ (i)}$$

in the first case.

Now insert the medium of dielectric constant κ between the condenser plates, the p.d. across the condenser being maintained equal to V. Then in theory and in practice (even with gases) an extremely narrow gap is left between the surface of the plates of the condenser and the boundaries of the medium. Owing to the polarization of the medium we get charges of surface density $\pm$ P at its boundaries, the charge due to polarization at the boundary near the positive plate of the condenser being negative, and vice versa (Fig. 37·5). The charge $-$ P induces an opposite charge $+$ P on the positive plate of the

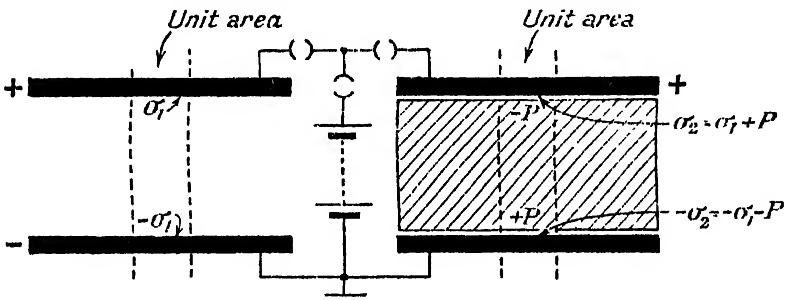


FIG. 37·5.—Electric Displacement.

condenser, so that the surface density of the electricity on the positive plate is now given by

$$\sigma_2 = \sigma_1 + P$$

But experimentally $\sigma_2 = \kappa\sigma_1$. Hence

$$\kappa\sigma_1 = \sigma_1 + P \qquad . \quad . \quad . \quad . \quad . \text{ (ii)}$$

From (i) we have

$$\kappa\frac{V}{4\pi d} = \frac{V}{4\pi d} + P$$

or

$$\kappa = 1 + 4\pi\frac{P}{\left(\frac{V}{d}\right)}$$

$$= 1 + 4\pi\left[\frac{\text{Polarization}}{\text{Applied field}}\right]$$

$$= 1 + 4\pi k \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \text{ (iii)}$$

where $k = P/E$, the *electric susceptibility* of the medium.

We also have, by multiplying (iii) throughout by E,

$$\kappa E = E + 4\pi P.$$

The quantity $E + 4\pi P$ is termed *dielectric displacement* or *electric induction* in the medium, and is denoted by D. Hence

$$D = \kappa E = E + 4\pi P \qquad . \quad . \quad . \quad . \text{ (iv)}$$

To see how D may be measured numerically, consider a needle-shaped cavity AB, Fig. 37·6 (*a*), in an insulator, the axis being drawn in the direction of the applied field. We assume the medium to be uniformly polarized, i.e. P is constant. If α is the cross-sectional area of this cavity, the charges developed at its ends are numerically equal to $P\alpha$. The electric intensity at O, a point near the centre of this cavity is due to the applied field of intensity E and the field due to the charges on the ends of the cavity. Since, numerically, α is small compared with the length of the cavity, the contribution to the electric intensity at O arising from the charges at the ends of the cavity must be zero. Hence the resultant electric intensity in such a cavity is E, the same as it would be if the insulator were replaced by a vacuum, i.e. $E = \frac{V_1 - V_2}{d}$.

When a cavity having the shape of a pill box is considered—cf. Fig. 37·6 (*b*)—the contribution to the electric intensity at a point near the centre of the cavity due to the charges on the ends of the cavity is no longer negligible. The point O may be regarded as one inside a parallel plate condenser the surface density of the electrification on the plates being numerically equal to P. The electric intensity due to such a distribution is $4\pi P$, so that the total electric intensity at O is $E + 4\pi P = D$. Hence D is numerically equal to the electric intensity in a cavity of the special type as defined above.

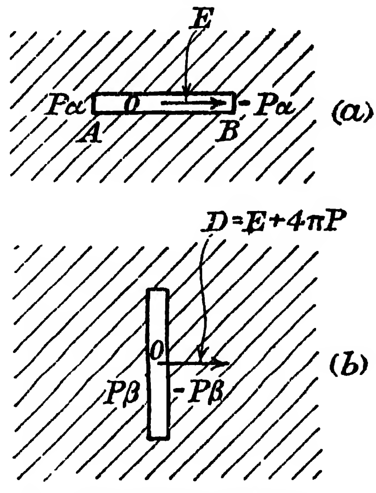


FIG. 37·6.—Electric Intensity and Electric Displacement.

The Law of Force for Dielectrics.—It has been shown that for condensers having the same size, but the space between the plates of one being filled with an insulator, the electric intensity is the same in each instance if the potential difference across the plates of each condenser is the same. But the charges on the plates of the condenser with a dielectric are κ times those on the corresponding plates of the other condenser. It therefore follows that the electric intensity in a dielectric would be $1/\kappa$-th that in a vacuum for equal charges. Hence, in general,

$$E = \frac{q}{\kappa r^2},$$

where q is a point charge in a medium whose dielectric coefficient is κ. The law of force between two point charges at distance r apart in a medium whose dielectric constant is κ is therefore

$$F = \frac{q_1 q_2}{\kappa r^2}.$$

Dielectric Constants of Gases : Variation with Temperature.—Let us suppose that each molecule of a gaseous dielectric has, on the average, an electric moment m, and that

$$m = \alpha E,$$

where E is the applied field and α a quantity whose nature is to be discovered. If there are n molecules per unit volume, then

$$nm = P = n\alpha E.$$

But $D = E + 4\pi P$

$$\therefore\ \kappa = \frac{D}{E} = 1 + 4\pi\frac{P}{E} = 1 + 4\pi n\alpha.$$

Hence, when κ is known, α may be determined. Quite naturally, therefore, we ask ourselves what does this quantity represent? Faraday imagined that α was a measure of the electric response of a molecule to an electric field—he imagined that the field caused a displacement of the two differently charged portions of a molecule so that a dipole was formed. The dipole ceased to exist when the field was removed.

Debye (1927) conceived the idea that a molecule may possess a definite electric moment even before the electric field is applied, i.e. the distribution of charge in such a molecule is asymmetrical. Debye termed them polar molecules, and the action of an electric field on them is twofold:

(i) it tends to align the molecules so that their electric axes are in the direction of the field;

(ii) it will cause a displacement of the electrons in each atom relative to the nucleus in that atom.

Debye therefore wrote

$$\alpha = \alpha_1 + \alpha_2,$$

where α_1 is the contribution caused by the application of an external field, and α_2 is the contribution due to the permanent electric moments

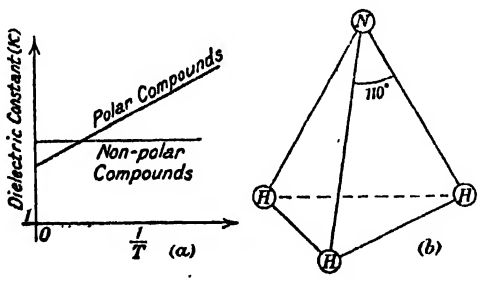


Fig. 37·7.
(a) Dielectric Constants of Gases and their Variation with Temperature.
(b) An NH_3-Molecule (polar).

of the molecules when a random orientation of their axes no longer exists. He argued that α_2 will depend on the temperature of the dielectric; as the temperature rises the molecules will be less liable to become orientated with their axes along the field. He showed that the dielectric constant of a gas whose molecules are permanent dipoles could be expressed by the formula

$$\kappa = a + \frac{b}{T},$$

where T is the absolute temperature and a and b are constants. For non-polar compounds κ is constant. If therefore κ is plotted against $1/T$, a straight line having a definite slope is obtained for gases whose molecules have a permanent electric moment; for non-polar compounds the straight line is parallel to the $1/T$ axis—c.f. Fig. 37·7 (a).

Ammonia, NH_3, and methane, CH_4, are examples of polar and non-polar molecules respectively. In the case of ammonia Debye has shown that the hydrogen atoms (really ions) are arranged at the base of a regular tetrahedron, the nitrogen ion occupying the apex—cf. Fig. 37·7 (*b*). The angle HNH is 110°.

It has just been shown that the electric intensity due to a point charge q at a point in a medium of dielectric constant κ and at distance r from the charge is given by

$$E = \frac{q}{\kappa r^2}$$

The electric induction or electric displacement at the same point is κE or q/r^2.

Gauss's Theorem.—Let us imagine that a closed surface is drawn in an electrostatic field. At each point on this surface the electric displacement has a value appropriate to that position. Since D is a vector it may be resolved into components along given directions. Let D_n denote the component of the displacement along the ***outward*** drawn normal at any point on the surface. Suppose that δS is a small

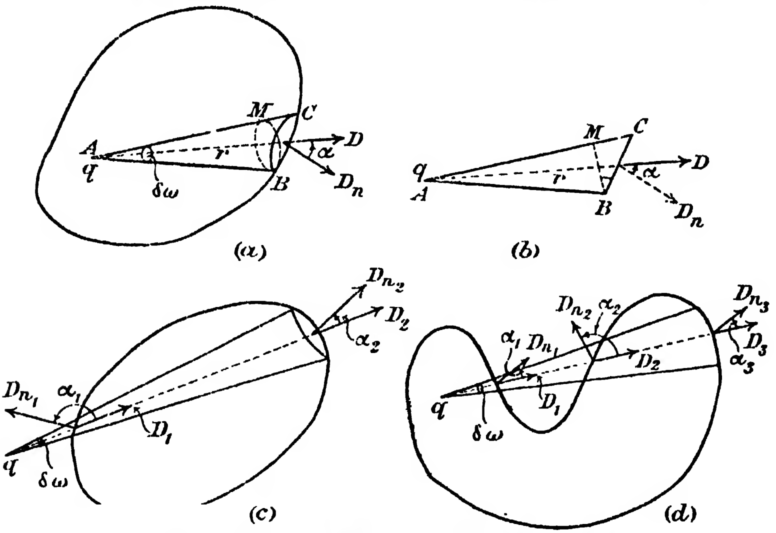


FIG. 37·8.—Gauss's Theorem in Electrostatics.

area across which the component D_n of the electric displacement may be considered constant. Then $D_n.\delta S$ is called the ***flux of electric induction*** or the ***normal induction*** cross the element δS.

Gauss's theorem for electrostatics states that ***the total flux of electric induction across any closed surface is 4π times the sum of the charges enclosed by that surface.*** Thus

$$\int D_n.dS = 4\pi q$$

where q is the charge enclosed.

PROOF.—Let q be the point charge at A inside a closed surface. Consider the flux of electric induction across a small element BC of

the closed surface—cf. Fig. 37·8 (*a*). It is equal to $D_n.\delta S$. Let α be the angle between the vector D and the outward drawn normal, and δS the area of the element. Then

$$D_n.\delta S = D.\delta S \cos \alpha = \frac{q}{r^2} \text{ (area BM)},$$

where MB is a section through B of the slender cone formed by joining all points on the periphery of δS to q, the section being at right angles to the axis of the cone—cf. Fig. 37·8 (*b*). But the area BM/r^2 is the measure of the solid angle at A—call it $\delta\omega$. Then the flux of induction across BC is $q\,.\,\delta\omega$. Hence the flux of electric induction across the closed surface is

$$\int q\,.\,d\omega = q\int d\omega = 4\pi q.$$

When the point charge lies outside the closed surface, every slender cone cutting the surface must do so at two places—cf. Fig. 37·8 (*c*). Let δS_1 and δS_2 be the areas of the closed surface intercepted by this elementary cone. Then the contribution from these two areas to the flux of electric induction is

$$\begin{aligned} D_{n_2}\,.\,\delta S_2 + D_{n_1}\,.\,\delta S_1 &= \frac{q}{{r_2}^2}\cos \alpha_2\,.\,\delta S_2 + \frac{q}{{r_1}^2}\cos \alpha_1\,.\,\delta S_1 \\ &= q\ \delta.\ \omega - \frac{q}{{r_1}^2}\cos(\pi - \alpha_1)\delta S_1 \\ &= q\,.\,\delta\omega - q\,.\,\delta\omega \\ &= 0 \end{aligned}$$

If the closed surface has a shape similar to that in Fig. 37·8 (*d*) and the point charge is within this surface, then a slender cone having its apex at the charge must intersect the surface an odd number of times. Let us assume that this number is three. Then the contribution to the flux of electric induction from the three elements of surface is

$$\begin{aligned} D_{n_1}\,.\,\delta S_1 &+ D_{n_2}\,.\,\delta S_2 + D_{n_3}\,.\,\delta S_3 \\ &= \frac{q}{{r_1}^2}\cos \alpha_1\,.\,\delta S_1 + \frac{q}{{r_2}^2}\cos \alpha_2\,.\,\delta S_2 + \frac{q}{{r_3}^2}\cos \alpha_3\,.\,\delta S_3 \\ &= q\,.\,\delta\omega - q\,.\,\delta\omega + q\,.\,\delta\omega \\ &= q\,.\,\delta\omega. \end{aligned}$$

Hence for the whole surface, the normal induction is given by

$$\int q\,.\,d\omega = 4\pi q.$$

Although the theorem has been established for a point charge, it applies to any distribution of charge, for this may be considered as a collection of point charges. Hence if $Q = \Sigma(q)$, is the total charge enclosed in a surface, Gauss's theorem states that

$$\int D_n dS = 4\pi Q.$$

Lines and Tubes of Electric Force and of Electric Induction.—We have seen that it is possible to draw continuous lines in an electric field in free space such that the tangent at any point to one of them indicates the direction of the electric intensity at that point. It is also possible to draw lines of force, for such is the name given to the

above lines, in a dielectric placed in an electrostatic field. The tangent to such a line in a dielectric then indicates the direction of the electric intensity in a small needle-shaped cavity having its centre at the point in question. It is also possible to draw other continuous curves in a dielectric—if these are such that the tangent at any point indicates the direction of the electric induction at that point, i.e. the force per unit charge on a small positive charge placed in a cavity having the shape of a pill box. Such lines are termed ***lines of induction.***

If any area is considered in an electric field and lines of force or lines of induction are drawn through every point on the contour of the above area, a tube of force or a tube of induction is obtained. In our study of dielectrics we shall find tubes of induction very helpful.

Tubes of Induction.—Let A, Fig. 37·9, be a point in an electric field, the electric induction or electric displacement at A, being D_1. Consider a small area δS_1 drawn round A so that it is normal to the direction of D_1. Construct the tube of induction having a cross-section δS_1 at A. Let δS_2 be the cross-section of the above tube at B, a point in the field where the electric displacement is D_2. Let us apply Gauss's theorem to this portion of the tube. The contribution to the flux of induction by the ends of the tubes is $D_2 . \delta S_2 - D_1 . \delta S_1$, since D_1 and D_2 are normal to the areas δS_1 and δS_2 respectively. The contribution from the curved sides of the tube is zero, since the normal component of the displacement is zero at all points on the curved portions of the tube. If there is no charge enclosed in the portion of the tube considered, then

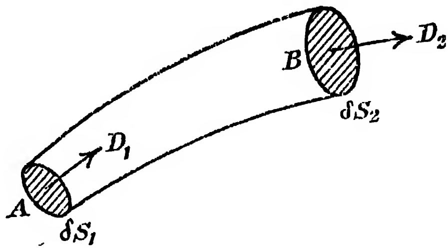


FIG. 37·9.—A Tube of Electric Induction.

$$D_2 . \delta S_2 - D_1 . \delta S_1 = 0.$$

Thus the electric displacement at any point is inversely proportional to the area of cross-section of the tube at that point.

The Electric Field inside a Hollow Charged Conductor.—Suppose that there is no electric charge inside a closed metal surface A, Fig. 37·10. Inside this surface construct any closed surface. Then the charge inside this surface is zero. By Gauss's theorem, the total flux of induction across this surface, being 4π times the charge enclosed, is zero, i.e.

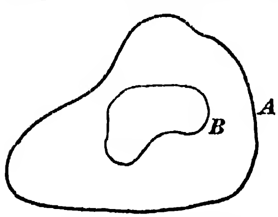


FIG. 37·10.—Electric Field inside a Closed Hollow Conductor.

$$\int D_n . dS = 0,$$

where D_n is the normal component of the electric displacement at a point on the surface. Since the above integral is zero for every closed surface which may be drawn inside A, it follows that D_n, and therefore D, must be zero at all points inside A. The electric intensity is also zero.

Coulomb's Theorem.—Let AB, Fig. 37·11 (*a*), be a small element of the surface of a charged conductor. Let σ be the density of the electricity on this area. Since the surface of a conductor is an equipotential surface, the electric induction must be at right angles to the surface. Consider the tube of induction whose cross-section at the surface is the element AB. Let MN be the cross-section of the tube a short distance away from the surface. Imagine that the tube is produced backwards and truncated by a plane PQ parallel to MN [PQ must lie inside the material of the conductor.] A section parallel to the axis of this cylinder is shown in Fig. 37·11 (*b*). The total normal induction over the surface of the cylinder thus obtained is due solely to the contribution from the end MN, since inside the conductor the electric intensity and therefore the electric induction is zero, and the normal component of the induction over the curved portions of the cylinder in the dielectric is zero at all points. The charge inside the cylinder is $\sigma \, . \, \delta S$. If D is the displacement at any point on MN, it is also the displacement at any point on AB when MN is sufficiently close to AB. The area of MN is then also δS. Hence by Gauss's theorem

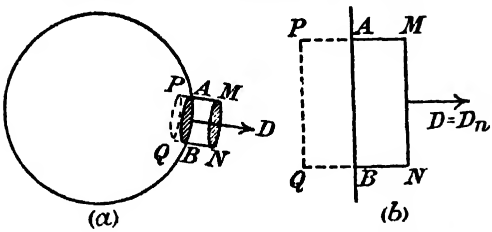


FIG. 37·11.—Coulomb's Theorem in Electrostatics.

$$D \, . \, \delta S = 4\pi \, . \, \sigma \delta S$$

or

$$D = 4\pi\sigma$$

The electric intensity near to the surface of a charged conductor is therefore given by

$$E = \frac{D}{\kappa} = \frac{4\pi\sigma}{\kappa}.$$

Electric Intensity and Electric Displacement due to a Uniformly Charged Sphere.

(i) ***At points outside the sphere.*** Let A, Fig. 37·12, be a point situated outside a sphere of radius a and carrying a charge q. Since the charge on the sphere is uniformly distributed, the electric displacement will be radial and equal at all points equidistant from O the centre of the sphere. Through A draw a sphere of radius OA, the centre being O. The electric displacement at all points on this sphere is D and is everywhere normal to the surface at the point considered. Hence the flux of induction across the sphere is

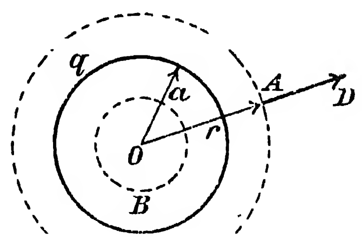


FIG. 37·12.—Electric Induction due to a Charged Sphere.

$$D \times (\text{area of surface of sphere of radius } r).$$
$$= 4\pi r^2 \, . \, D.$$

But by Gauss's theorem this is $4\pi q$. Hence

$$D = \frac{q}{\ }.$$

The electric intensity is therefore   At points outside the sphere the electric displacement and intensity are respectively the same as if the charge were concentrated at the centre of the sphere.

(ii) ***At points inside the sphere.*** If B is such a point and a sphere, with centre O and radius OB, is constructed, the electric displacement must again be radial at all points on the sphere. The flux of induction across this sphere is zero, for by Gauss's theorem it is 4π times the charge inside the sphere, and the charge inside the sphere is zero. Hence, inside the sphere, D, and therefore E, are zero.

Uniformly Electrified Infinite Flat Plate.—Let σ be the density of the electricity on ***each*** side of the plate. By symmetry, the electric displacement must everywhere be normal to the plate and have a constant value at all points in a plane parallel to it. Under these conditions the area of cross-section of any tube of induction remains constant. Consider the portion of such a tube shown in Fig. 37·13. Let D_1 and D_2 be the values of the electric displacement at the lower and upper ends of this cylinder. Let the areas of these ends be s. It is not necessary for these areas to be small since, by symmetry, D_1 and D_2 are constant at all points in the respective planes parallel to the given plate. The flux of induction across the surface of this cylinder is

$$D_2 s - D_1 s,$$

the contribution from the curved surfaces being zero. The above is zero, by Gauss's theorem. Hence

$$D_1 = D_2,$$

or the electric displacement is constant.

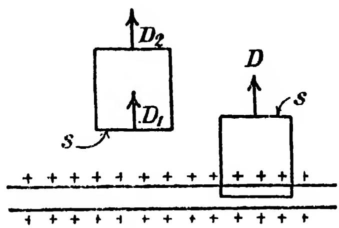


FIG. 37·13.—Electric Induction due to a Charged Infinite Plate.

To determine the value of the displacement consider the portion of a tube of induction indicated on the right of the diagram. This tube originates on a charge σs. Let the above cylinder be produced backwards so that it may be truncated inside the conductor. There can be no tubes of induction inside the charged plate, but that does not prevent us from drawing a portion of a closed surface inside the conductor. The only contribution to the flux of induction arises from the end of the tube in the dielectric, and amounts to $D . s$. By Gauss's theorem this is $4\pi\sigma s$, so that $D = 4\pi\sigma$. In using this formula it must be remembered that σ is the surface density of the electricity on one side of the plane only.

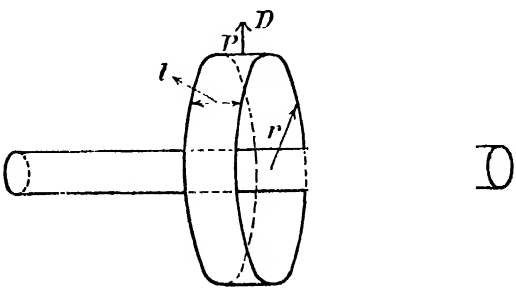


FIG. 37·14.—Electric Intensity due to an Electrified Infinite Cylinder.

Uniformly Electrified Infinite Cylinder.—Let P, Fig. 37·14, be a point at a distance r

from the axis of an infinite cylinder, the charge of electricity per unit length being λ. By symmetry D will be radial and have the same value at all points equidistant from the axis of the cylinder. Through P describe a cylinder coaxial with the charged one, and construct two planes at distance l apart and normal to the axis of the cylinder to form a closed surface. The charge enclosed by this surface is λl. The plane ends contribute nothing to the flux of induction across the closed surface considered. The flux across the curved surface of the cylinder is $D \cdot 2\pi rl$. By Gauss's theorem this is $4\pi\lambda l$. Hence

$$D = \frac{2\lambda}{r},$$

and

$$E = \frac{2\lambda}{\kappa r},$$

where κ is the dielectric coefficient for the medium surrounding the charged cylinder.

Inside the cylinder the electric displacement is zero.

The Mechanical Stress at the Surface of a Charged Conductor.—Let A and B, Fig. 37·15, be two points close to the surface of a charged conductor, one outside and the other inside. Let κ be the specific inductive capacity of the medium. The problem before us is to investigate the force on the charge residing on unit area of the surface of the conductor. This is determined by E, the electric intensity at A, which may be regarded as the resultant intensity due to the electricity on the surface of the conductor in its immediate neighbourhood and to the rest of the distribution. Let these contributions be E_1 and E_2 respectively. Then

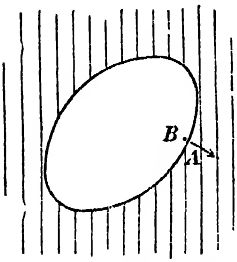


FIG. 37·15.—Mechanical Stress at the surface of a Charged Conductor.

$$E = E_1 + E_2.$$

At B the electric intensity is zero. The contribution to this is $-E_1$ from the electricity on the surface, and E_2 from the remainder of the distribution. Hence, $-E_1 + E_2 = 0$, or $E_1 = E_2 = E/2$.

Consider a small area δS of the surface of the conductor. Then the charge on it is $\sigma . \delta S$. Now the force acting on the charge $\sigma . \delta S$ is caused by the electric intensity at the point considered due to the rest of the electricity on the conductor, i.e. the force is $\sigma . \delta S . E_2$. $= \frac{1}{2}\sigma . \delta S . E$.

The force per unit area is therefore $\frac{1}{2}\sigma E$. But $D = 4\pi\sigma$, so that the expression for the above force per unit area, or surface stress, becomes

$$\frac{DE}{8\pi} = \frac{\kappa E^2}{8\pi} = \frac{2\pi\sigma^2}{\kappa}.$$

The sign of this stress is independent of that of the charge and it is always directed outwards.

Experiment.—The existence of the above stress may be shown by producing a soap bubble at the end of an insulated metal tube. When the tube and therefore the soap bubble is connected to a source of high potential the bubble expands until the reduction in pressure inside the bubble compensates for the surface stress arising from the charge on the bubble.

Energy in an Electrostatic Field.—Consider an element δS of the surface of a charged conductor—Fig. 37·16. Let σ be the surface density of the electricity. Then the force acting on the portion of the surface considered is $\frac{DE}{8\pi}.\delta S$. Suppose that the surface is moved backwards a distance δx. Then the work done by the external agency causing the motion is $\frac{DE}{8\pi}.\delta S.\delta x$. But the field has been increased in volume by an amount $\delta S.\delta x$. The work done in creating a field is stored as energy in the medium. The energy associated with unit volume of the field is therefore

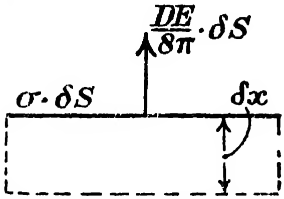


FIG. 37·16.—Energy in an Electrostatic Field.

$$\frac{DE}{8\pi}$$

Stresses in an Electrostatic Field.—The mechanical force acting on unit area of an electrified conductor is $\frac{DE}{8\pi}$ or $\frac{2\pi\sigma^2}{\kappa}$. This stress may be regarded as the effect of a tension along the tubes of induction in the field. We assume that the tension per unit cross-sectional area of a tube at any point in the field is given by the above expression,

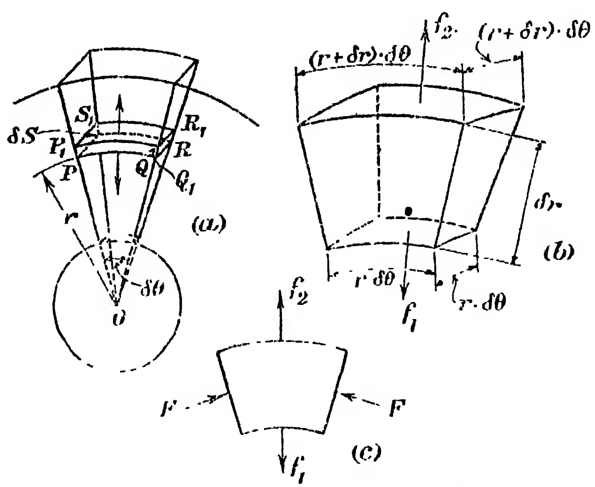


FIG. 37·17.—Stresses in an Electrostatic Field.

subject to the condition that D and E are the values appropriate to that point. Let us see whether or not it is possible for a tube of induction to be in equilibrium under the action of these forces alone.

As a special case consider a tube of induction in the region between two concentric spheres—Fig. 37·17 (a). Suppose that the cross-section of the tube at right angles to its axis at any point is a square, and that opposite sides of the tube are inclined to each other at an angle $\delta\theta$. Let q be the charge on the inner sphere. Consider that portion of

the above tube bounded by the surfaces of spheres of radii r and $(r + \delta r)$. This element of the tube is shown enlarged in Fig. 37·17 (*b*). Let f_1 and f_2 be the forces acting on the opposite curved ends of this element due to the tension along the tube. The tension across the face of radius r is $\frac{\mathrm{DE}}{8\pi}$ per unit area. Hence

$$f_1 = \frac{\mathrm{DE}}{8\pi} \cdot (r\delta\theta)^2 = \frac{\kappa}{8\pi} \cdot \mathrm{E}^2 \cdot (r \cdot \delta\theta)^2 = \frac{\kappa}{8\pi}\left(\frac{q}{\kappa r^2}\right)^2 \cdot r^2 \cdot \delta\theta^2 = \frac{q^2\delta\theta^2}{8\pi\kappa r^2}$$

since $\mathrm{D} = \kappa\mathrm{E} = q/r^2$. Similarly

$$f_2 = \frac{q^2\delta\theta^2}{8\pi\kappa(r + \delta r)^2}$$

Hence
$$f_1 - f_2 = \frac{q^2\delta\theta^2}{8\pi\kappa}\left[\frac{1}{r^2} - \frac{1}{(r + \delta r)^2}\right]$$
$$= \frac{q^2\delta\theta^2}{8\pi\kappa} \cdot \frac{2 \cdot \delta r}{r}, \text{ [neglecting } \delta r^2\text{]}.$$

Thus there is a resultant force directed *inwards* and the tube cannot therefore be in equilibrium under the action of the above tensions. Suppose, however, that P is a pressure (i.e. a force per unit area) acting on each of the four flat sides of the element. Then the force on each side is $\mathrm{P}r\delta\theta \cdot \delta r$. The component of this force along the axis of the tube is $\mathrm{P}r\delta\theta \cdot \delta r \cdot \frac{\delta\theta}{2}$. The resultant force arising from the pressure acting on the flat sides of the element is therefore

$$4 \cdot \frac{\mathrm{P}r\delta\theta^2 \cdot \delta r}{2} = 2\mathrm{P}r\delta r \cdot \delta\theta^2$$

along the axis of the tube and away from the centre of the spheres. For equilibrium this must equal $f_1 - f_2$, i.e.

$$\mathrm{P} = \frac{q^2}{8\pi\kappa r^4} = \frac{1}{8\pi} \cdot \frac{q}{r^2} \cdot \frac{q}{\kappa r^2} = \frac{\kappa\mathrm{E}^2}{8\pi}$$
$$= \frac{\mathrm{DE}}{8\pi}.$$

Hence for a tube of induction to be in equilibrium there must be a tensile stress $\frac{\mathrm{DE}}{8\pi}$ along the tube and a compressive stress $\frac{\mathrm{DE}}{8\pi}$ normal to the sides of the tube.

The existence of this pressure is shown by the following experiment due to QUINCKE:—Two large metal plates are insulated from each other by pieces of glass and the whole immersed in paraffin oil—Fig. 37·18. By blowing through a tube containing calcium chloride a bubble is produced at B. One of the plates is connected to a Wimshurst machine [cf. p. 736] and charged. To determine the lateral pressure between the tubes of induction in the liquid and in the air bubble let us

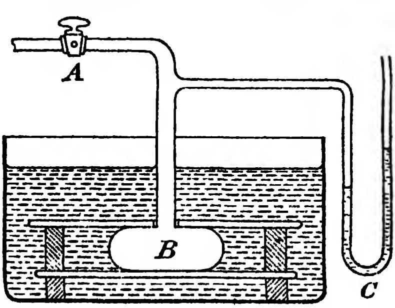


FIG. 37·18.—Quincke's Experiment.

consider the condensers formed by unit areas of the metal plates and the intervening dielectric, (*a*) in the bubble and (*b*) in the liquid. If d is the distance apart of the plates, the capacities of the two unit condensers are $\frac{1}{4\pi d}$ and $\frac{\kappa}{4\pi d}$ respectively. If V is the difference in potential between the plates, the charges on the plates of the unit condensers are $\frac{V}{4\pi d}$ and $\frac{\kappa V}{4\pi d}$ which are denoted by σ_a and σ_l respectively. Hence, in the liquid the compressive stress is

$$\frac{DE}{8\pi} = \frac{\kappa V^2}{8\pi d^2}, \text{ since } D = \kappa E \text{ and } E = \frac{V}{d}.$$

In the bubble itself the compressive stress between adjacent tubes of induction is $\frac{V^2}{8\pi d^2}$. The repulsion of the tubes of induction is, therefore, greater in the oil than in air so that the bubble contracts and the gauge C indicates an increase in pressure.

The Capacitance of a Concentric Spherical Condenser.—Let O —Fig. 37·19 (*a*)—be the centre of two concentric spheres of radii a and b respectively. Let q be the charge on the inner sphere; then $-q$ is the induced charge on the outer sphere when this is earthed. Let κ be the permittivity of the medium between the spheres. Let E be the electric intensity at P, a point in the dielectric at distance r from O. Then

$$E = \frac{q}{\kappa r^2}$$

Let V be the potential of the inner sphere, that of the outer one being zero. Then

V = the work done against the field per unit positive charge in bringing a small positive charge from the outer sphere to the inner sphere $= -\int_b^a \frac{q}{\kappa r^2}dr$. [Cf. p. 684 for the reason why the negative sign occurs.]

$$\therefore V = \frac{q}{\kappa}\left[\frac{1}{r}\right]_b^a = \frac{q}{\kappa}\left(\frac{1}{a} - \frac{1}{b}\right).$$

The capacitance is therefore $\frac{q}{V} = 1\Big/\frac{1}{\kappa}\left(\frac{1}{a} - \frac{1}{b}\right) = \frac{\kappa ab}{b - a}$.

An interesting problem is presented when the inner sphere is earthed and the outer spherical shell of the previous system is insulated and charged—Fig. 37·19 (*b*). Let c be the radius of the outer surface of the spherical shell. If q is the charge on the inner surface of the shell, $-q$ is the induced charge on the sphere when this is earthed; let Q be the charge on the outer surface of the shell. The electric intensity in the dielectric is due solely to the charge $-q$ on the inner sphere: it is $-\frac{q}{\kappa r^2}$.

The potential V of the shell is constant throughout and is given by

O − V = work done per unit charge in taking a small positive charge from the surface $r = b$ to the surface $r = a$.

$$= \int_b^a - \frac{q}{\kappa r^2}(-dr)$$

$$= \frac{q}{\kappa}\left(-\frac{1}{a} + \frac{1}{b}\right)$$

$$\therefore\ V = \frac{q}{\kappa}\left(\frac{1}{a} - \frac{1}{b}\right).$$

But the potential of the outer shell is also given by

$$\frac{Q}{c} + \frac{q}{c} - \frac{q}{c} = \frac{Q}{c},$$

since charges on spherical conducting surfaces act at external points and points on their own surfaces as if they were concentrated at the centre of the sphere.

$$\therefore \frac{q}{\kappa}\left(\frac{1}{a} - \frac{1}{b}\right) = \frac{Q}{c} = V.$$

The total charge on the shell is therefore $Q + q$, so that the capacitance of the system is given by

$$C = \frac{Q+q}{V} = \frac{q\dfrac{c(b-a)}{\kappa ab} + q}{V} = \frac{q\left[\dfrac{c(b-a)}{\kappa ab} + 1\right]}{q\dfrac{(b-a)}{\kappa ab}} = c + \frac{\kappa ab}{b-a}$$

This result can be written down at once if we regard the condenser as consisting of two parts, of capacitances $\dfrac{\kappa ab}{b-a}$ and c, arranged in parallel.

Alternative Method.—When air is the dielectric let the charges on the surfaces be q_a, q_b and q_c respectively. Then, remembering that the potential inside a conductor due to its own charge is the same as that of the conductor itself, we have, as the potential of the inner sphere

$$-\frac{q_a}{a} + \frac{q_b}{b} + \frac{q_c}{c} = o, \quad . \quad . \quad . \quad . \quad (i)$$

since it is earthed. Moreover, the potential of the inner surface of the shell must be the same as that of its outer surface. Hence,

$$-\frac{q_a}{b} + \frac{q_b}{b} + \frac{q_c}{c} = \frac{q_c}{c} + \frac{q_b}{c} - \frac{q_a}{c} \quad . \quad . \quad . \quad (ii)$$

$$\therefore\ q_a = q_b = q(\text{say}) \quad . \quad . \quad . \quad . \quad (iii)$$

$\therefore$ From (i), $$q\left[\frac{1}{b} - \frac{1}{a}\right] + \frac{q_c}{c} = o \quad . \quad . \quad . \quad . \quad (iv)$$

The potential of the shell is, from (ii),

$$\frac{q_c}{c}.$$

This is equal to the total charge on the shell, divided by the capacitance of the condenser.

$$\therefore\ \text{Capacitance} = \frac{q + q_c}{\dfrac{q_c}{c}} = c\left(1 + \frac{q}{q_c}\right),$$

$$= c + \frac{ab}{b-a}. \qquad \text{(from iv)}$$

The Capacitance per Unit Length of a Long Coaxial Cylindrical Condenser.—Let Fig. 37·19 (*a*) represent the cross-section of a long coaxial cylindrical condenser, the outer shell being earthed while the

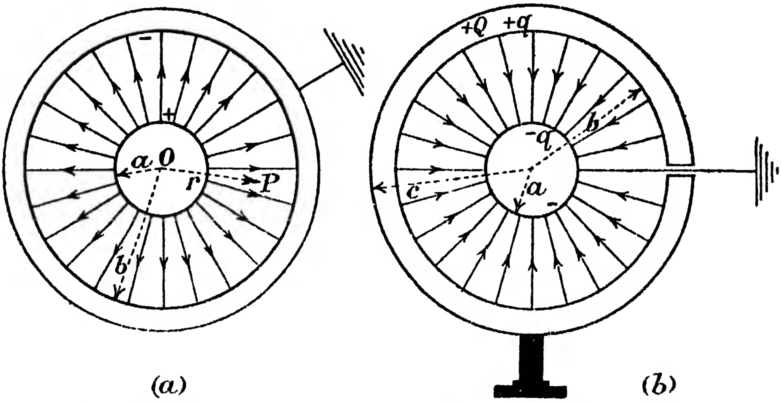


FIG. 37·19.—Capacitance of a Concentric Spherical Condenser.

inner cylinder is insulated and carries a charge λ per unit length. The intensity at P is $2\lambda/\kappa r$. Hence

$$V_b = -\int_b^a \frac{2\lambda}{\kappa r}dr = \frac{2\lambda}{\kappa}[\log_e b - \log_e a].$$

$$\therefore \text{ Capacitance per unit length} = \frac{\kappa}{2[\log_e b - \log_e a]},$$

$$= \frac{\kappa}{2\log_e \dfrac{b}{a}}.$$

A Variable Cylindrical Air Condenser.—To the ends of a wooden base there are attached two ebonite uprights carrying a long ebonite rod AB, Fig. 37·20 (*a*). This rod supports a brass tube M, the outer radius of which is *a*. N is a coaxial brass cylinder carried on two metal supports to one of which is fixed a spring S making contact with an

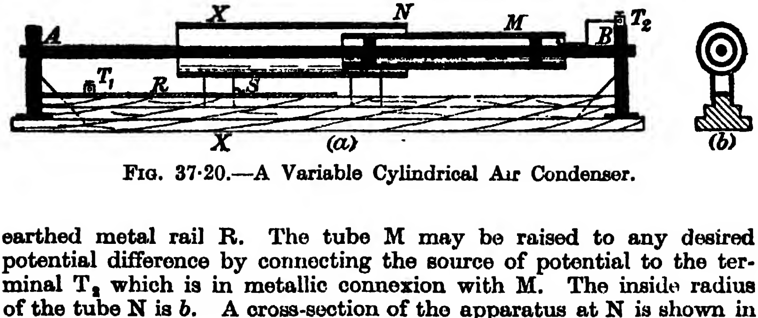


FIG. 37·20.—A Variable Cylindrical Air Condenser.

earthed metal rail R. The tube M may be raised to any desired potential difference by connecting the source of potential to the terminal T_2 which is in metallic connexion with M. The inside radius of the tube N is *b*. A cross-section of the apparatus at N is shown in Fig. 37·20 (*b*).

The absolute capacitance of such a condenser is unknown but by

moving N a distance l to the right or left the capacitance may be respectively increased or decreased by an amount

$$\frac{l}{2 \log_e \frac{b}{a}}. \qquad \text{[e.s.u. capacitance.]}$$

Cavendish's Experimental Verification of the Inverse Square Law in Electrostatics.—A metal globe was suspended from an insulating support. An insulated spherical shell, concentric with the globe, was formed by fastening two metal hemispheres by glass rods to two wooden frames hinged to an axis so that the hemispheres could be placed in the desired position.

The globe could be put into metallic communication with the hemispheres by means of a short wire insulated by a silk thread, so that it was capable of being removed without discharging the apparatus. Metallic connexion between the globe and hemispheres having been made, both were connected to a Leyden jar whose potential had been measured by an electrometer. The wire was withdrawn, the hemispheres removed and discharged, and the electrical condition of the globe tested by means of a pith ball electrometer, which at that time (1773) was regarded as the most delicate electroscope. No indication of any charge on the globe was detected.

Cavendish then communicated to the globe a known fraction of the charge originally given to the spherical condenser and tested the electrical state of the globe. In this way he found that the charge on the globe in the first experiment was less than $\frac{1}{60}$ that supplied to the condenser, for greater charges than this were detected by the electrometer.

He then calculated the fraction of the charge which would have remained on the globe if the law of repulsion between like charges differed by a small quantity from that of the inverse square. If this difference were $\frac{1}{50}$ the fractional charge on the globe would have amounted to $\frac{1}{57}$ of that on the condenser. Such a charge would have been detected with his apparatus.

Maxwell's Experimental Verification of the Inverse Square Law in Electrostatics.—The metal hemispheres were supported on an insulating stand, the inner sphere being held in position by means of an ebonite ring A, Fig. 37·21. In this way the insulating support for the inner sphere was never exposed to the action of any electric field, and therefore never received any charge which might have been a disturbing factor. Instead of removing the hemispheres before testing the globe for electricity, they were allowed to remain in position. In this way the inner sphere was protected from all external electric fields, an advantage far outweighing the disadvantage due to the fact that the effect of a given charge on the inner sphere was not so great as if the hemispheres had been removed, i.e. the capacity of the electrometer and its connexions was increased.

The short wire, B, making metallic communication between the inner and outer spheres was attached to a small metal disc covering the aperture in the shell. When the lid and wire were raised by means of a silk thread, the electrode attached to the electrometer (Kelvin's) could be brought into contact with the inner sphere. The case of the electrometer, one pair of quadrants, and the exploring electrode, T, Fig. 37·21 (c), were all connected to earth until the shell had been discharged.

To estimate the original charge of the shell, a small brass sphere was placed on an insulating stand at a considerable distance from the rest of the apparatus.

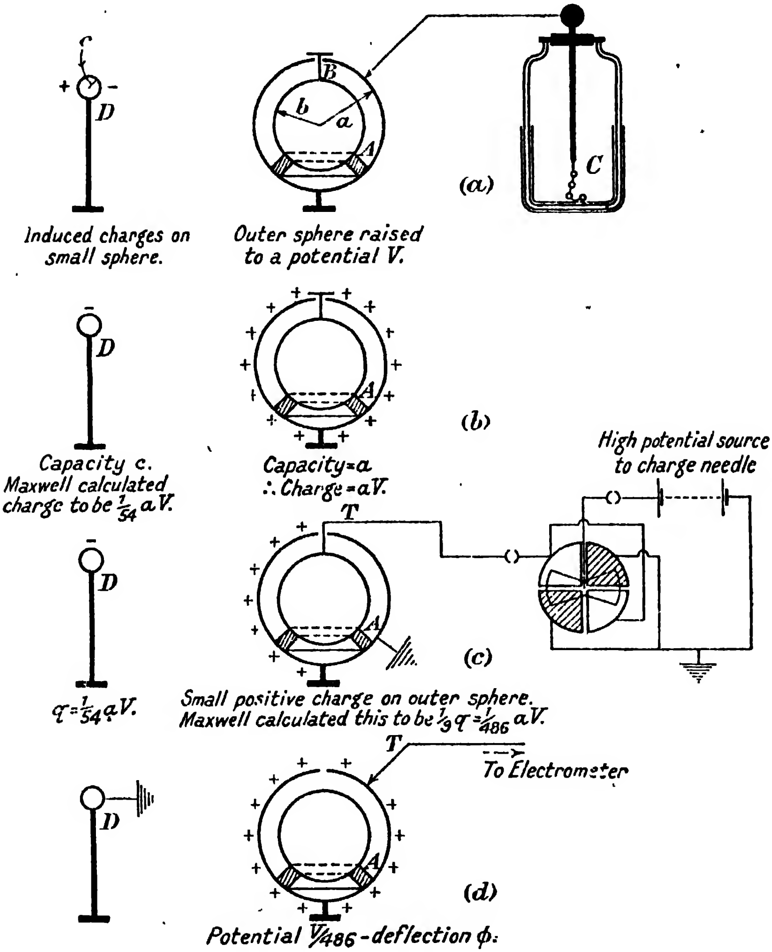


FIG. 37·21.—Maxwell's Investigation of the Validity of the Inverse Square Law in Electrostatics.

The following operations were carried out:

(i) The shell was charged by communication with a Leyden jar, C, Fig. 37·21 (*a*), the wire B making connexion between the inner and outer spheres.

(ii) The small brass ball, D, was earthed so that it received a charge by induction. It was then insulated—cf. Fig. 37·21 (*b*).

(iii) The communicating wire was withdrawn.

(iv) The outer shell was earthed.

(v) The testing electrode, T, was brought into contact with the inner sphere—Fig. 37·21 (*c*).

'Not the slightest effect on the electrometer could be observed,' writes Maxwell.

To test the sensitivity of the apparatus the shell was disconnected from earth, and since it had been under the influence of a negative charge on the small sphere it had acquired a positive charge. The small ball was then discharged and, the testing electrode attached to the electrometer being in contact with the outer sphere, there was a deflection ϕ—Fig. 37·21 (*d*).

The negative charge on the ball was about $\frac{1}{54}$ of the original charge on the shell, and the positive charge induced on the shell was about $\frac{1}{9}$ of that on the ball. Hence the potential of the shell, as indicated by the electrometer, was about $\frac{1}{486}$ of its original potential.

Maxwell then calculated that if the repulsion had followed the law $r^{\beta-2}$, the potential of the inner sphere would have been $-0{\cdot}1478\beta V$ where V is the potential of the shell. Suppose that $\pm\sigma$ is the smallest deflexion of the electrometer needle which could be detected. Then

$$0{\cdot}1478\beta V < \pm k\sigma \qquad [k = \text{conversion factor}]$$

and

$$\frac{V}{486} = k\phi,$$

$$\therefore\ 72\beta < \pm\frac{\sigma}{\phi}.$$

Now ϕ was certainly 300 times greater than σ,

$$\therefore\ \beta < \pm\frac{1}{21600}.$$

It was in order to estimate the potential of the inner sphere if the value of β were not zero that both Cavendish and Maxwell used spherical condensers, but in any closed conductor the electric intensity is zero.

Parallel Plate Condenser.—Let us suppose that A and B, Fig. 37·22, are the plates of such a condenser, the distance between the plates being small compared with the linear dimensions of the plates, so that we may be justified in regarding the lines of induction as normal to the plates over their central regions. Let V_2 and V_1 be the potentials of the plates, the densities of the charges on the plates being σ and $-\sigma$ respectively. Let KL be a small element of area δs parallel to the surface of either plate. Construct the closed surface having δs for one base and straight lines through each point on the periphery of this base forming its curved surface. Let this element be truncated by a plane, MN, parallel to KL and inside the plate A. Then the charge enclosed by this Gaussian surface is $+\sigma\delta s$. Let D be the electric displacement at any point on the base KL. Then the direction of this displacement is normal to KL. The contribution to the flux of induction is zero across the curved surface of the element

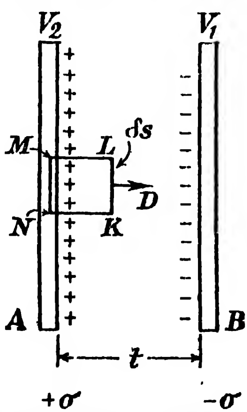


FIG. 37·22.—A Parallel Plate Condenser.

and also across MN, since this is a surface inside the conductor. Hence, by Gauss's theorem,

$$D \,.\, \delta s = 4\pi \,.\, \sigma\delta s$$
$$\therefore\ D = 4\pi\sigma.$$

If κ is the dielectric coefficient of the medium between the plates of the condenser, the electric intensity is given by

$$E = \frac{4\pi\sigma}{\kappa}.$$

If t is the distance between the two plates.

$$V_2 - V_1 = E \,.\, t = \frac{4\pi\sigma t}{\kappa}.$$

Since σ is the charge per unit area of the positive plate of the condenser, the above expression shows that the capacitance per unit area of the above condenser is $\frac{\kappa}{4\pi t}$.

If the dielectric has a constant width d, where $d < t$, then

$$V_2 - V_1 = E_1(t - d) + E_2 d,$$

where E_1 and E_2 are respectively the electric intensities in the air and in the dielectric.

If σ is the surface density on the positive plate under these conditions, $E_1 = 4\pi\sigma$ and $E_2 = 4\pi\sigma/\kappa$. Hence

$$\begin{aligned} V_2 - V_1 &= 4\pi\sigma(t - d) + \frac{4\pi\sigma}{\kappa}d \\ &= 4\pi\sigma\left[t - d + \frac{d}{\kappa}\right]. \end{aligned}$$

The capacitance per unit area is therefore

$$\frac{1}{4\pi\left[t - d\left(1 - \frac{1}{\kappa}\right)\right]}.$$

EXAMPLES XXXVII

1.—State Gauss's Theorem in electrostatics. Apply it, (*a*) to show that the product field-strength $\times$ cross-section, along a tube of force in air and containing no charge is constant; (*b*) to obtain an expression for the strength of an electric field just outside a charged conductor at a place where the surface density of the charge is σ.

2.—A potential difference of 2000 volts exists between two large parallel plates in air, at a distance apart of 1 cm. Calculate the pull on unit area of each plate. What would be the effect on this pull if the space between the plates were filled with an oil of dielectric constant 3·5?

3.—The electrostatic potential at a point in air near the earth's surface increases with height at a rate of 100 volt.metre.$^{-1}$ Calculate the charge per square metre on the earth. Calculate also the resulting mechanical stress at the earth's surface in the same locality.

4.—Derive an expression for the capacitance per unit area of a large parallel plate condenser which has half the distance between its plates occupied by a slab of material of dielectric constant κ, and the remaining half by air.

5.—Assuming that the forces in an electrostatic field can be ascribed to a system of stresses in the dielectric medium, obtain an expression for the tensions along the lines of force. What other stress is necessary for equilibrium? How has its existence been demonstrated and its value verified?

Calculate the pull per unit area on the surface of a charged conductor at a place where the surface density of the charge is 10^{-9} coulomb cm.$^{-2}$, and the surrounding dielectric has a constant equal to three times that of empty space.

6.—Calculate a value for the radius of a water drop which, carrying a negative charge equal to that of an electron ($-4{\cdot}77 \times 10^{-10}$ e.s.u.) floats in the earth's electric field when the vertical intensity is 150 volt. metre.$^{-1}$ Is the general direction of the field upwards or downwards in this case?

CHAPTER XXXVIII

ELECTROSTATIC INSTRUMENTS

ELECTROSTATIC MEASURING INSTRUMENTS

The Bifilar Electrometer.—A section of this instrument is indicated in Fig. 38·1. AA is a loop of thin platinum wire stretched between a metal rod C [insulated by amber from the case of the instrument] and a circular piece of quartz D. BB are wires fixed to the walls which are earthed. When C is connected to a source of potential, electrical attraction causes the wire loop A to expand sideways. The displacement is measured with the aid of a microscope and is a measure of the potential difference between A and the earth. Potential differences from 30 to 300 volts may be measured in this way. [N.B.—The volt is *not* the electrostatic unit of potential difference. Unit potential difference on the c.g.s. electrostatic system of units is equivalent to a potential difference of 300 volts.]

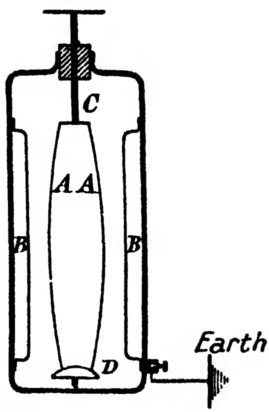


FIG. 38·1.—Bifilar Electrometer.

The Attracted Disc Electrometer.—This electrometer consists essentially of a guard-ring condenser and a balance, the underlying principle being that the mechanical pull on the movable plate of a condenser is balanced against the gravitational pull on a known mass. A, Fig. 38·2, is the lower plate of the condenser. It is supported by an insulated screw and may be raised to any desired potential. The upper plate of the condenser consists of a central circular section B, surrounded by a wide concentric ring C. This is the so-called guard-ring. B is supported from one arm of a balance as shown. The clearance between B and C is sufficient for B to move freely and yet not sufficient to disturb the homogeneity of the field in the central region of the condenser. C, and the support for the beam of the balance, are earthed, i.e. B is per-

manently earthed also. The balance is first adjusted so that B lies flush in the plane containing C. The beam is then just in contact with the stop K. This adjustment may be effected by adding sand to the balance pan D.

Let V be the potential of the lower plate at a distance t from the upper one and m be the additional mass required in D to restore

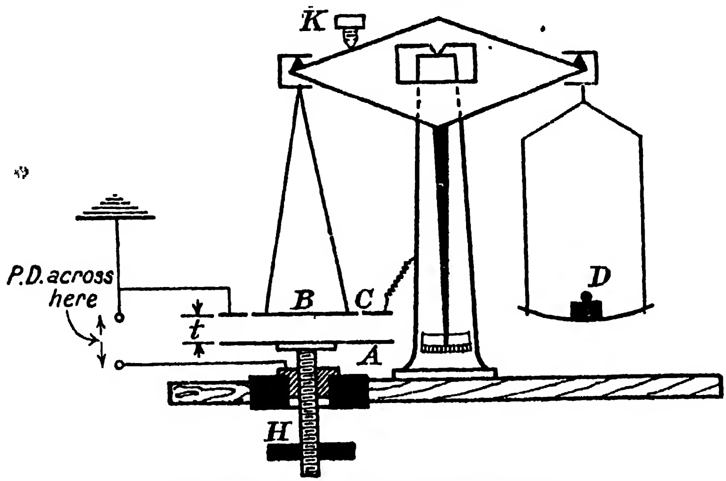


FIG. 38·2.—Attracted Disc Electrometer.

equilibrium, i.e. to make B and C coplanar again. Let σ be the numerical value of the density of the charges on the central portions of the condenser plates. If S is the area of the plate B (strictly the mean area of the aperture in C and the plate B), the total pull, F, on it is $2\pi\sigma^2 S$. But

$$\frac{V}{t} = \text{electric intensity} = 4\pi\sigma,$$

$$\therefore\ V = 2t\sqrt{\frac{2\pi mg}{S}}.$$

Since all the quantities on the right-hand side of the above equation are known, or measurable, V may be calculated.

In actual practice it is found difficult to measure t accurately, so that the following modified procedure is adopted. Let A be connected to a constant source of potential, V, the plates being at a distance t apart. Then

$$V = 2t\sqrt{\frac{2\pi mg}{S}}.$$

Now let the potential to be measured, say v, be connected in series with V, the total potential being $V + v$. Let A be moved,

by means of the screw H, through a vertical distance h until the balance is again equilibrated.

$$V + v = 2(t + h)\sqrt{\frac{2\pi mg}{S}}.$$

From these we have

$$v = 2h\sqrt{\frac{2\pi mg}{S}}.$$

The Quadrant Electrometer.—This instrument enables us to compare potential differences more accurately than can be done with gold-leaf electroscopes. LORD KELVIN made the first reliable quadrant electrometer, but the form chiefly used to-day is due to DOLEZALEK [cf. Fig. 38·3 (a)]. It consists essentially of a cylindrical box divided into quadrants, one of which is shown in Fig. 38·3 (b). Diagonally opposite quadrants are connected by thin wires and an aluminium needle is suspended symmetrically in a horizontal plane

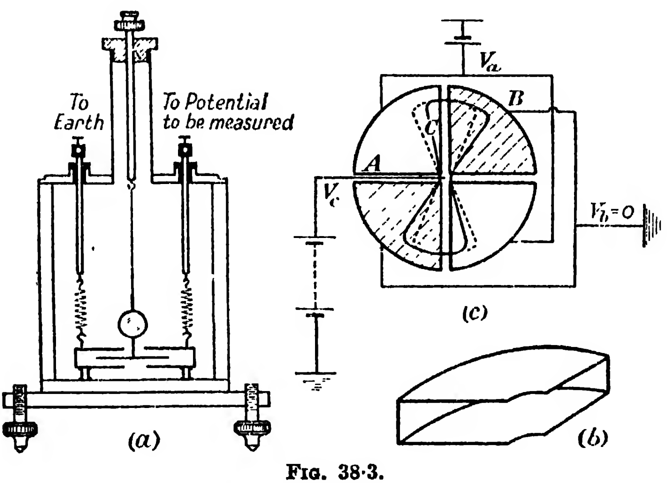


FIG. 38·3.

(a) Dolezalek Quadrant Electrometer.
(b) A quadrant.
(c) Diagrammatic representation of a quadrant electrometer and its connexions.

between them. The needle is suspended by a fine phosphor bronze wire which is raised to about 100 volts by being connected to one terminal of a battery, the other terminal being earthed—cf. Fig. 38·3 (c). Each quadrant is supported on a quartz pillar which should never be touched by the hand if the insulation is to remain unimpaired. Communication to the quadrants is made by means of metal rods and springs passing through the brass case surrounding the instrument but insulated from it.

In some of these instruments the suspension consists of a quartz fibre which is chosen on account of its constant elastic properties. The needle is then charged by touching it with a charged rod. The insulation resistance of the quartz is so high that the needle does not lose its charge for a considerable time. The base of the instrument is fitted with screws so that it may be levelled and it is advisable to surround the entire instrument with an earthed piece of gauze to protect the quadrants and needle from stray electric fields. A mirror is rigidly attached to the needle so that small angular displacements of the needle may be measured.

The principle underlying this instrument is that when there is a difference of potential between the two pairs of quadrants, the needle, having a positive charge, moves away from the quadrants with the higher potential. The energy of the needle is spent in doing work in twisting the fibre. For small potential differences the deflexion of the needle is proportional to the potential difference.

If the instrument is not exceptionally sensitive—say it gives a deflexion of 10 cm. on a scale 1 metre away for a p.d. of one volt—then it may be used to compare the e.m.fs. of two cells by first earthing the quadrants and determining the zero of the instrument. One pole of one of the cells is then earthed and the other connected to one pair of quadrants [disconnected from earth] and the deflexion observed. The second cell is then examined in the same way. The ratio of the E.M.F.'s of the cells is the ratio of the deflexions of the needle.

To Determine the Capacitance of an Electrometer and its Connexions.—Let E, Fig. 38·4, be the electrometer whose needle is maintained at a high potential (100 V.) by means of a battery A. It will be supposed that a fixed air condenser, C', is permanently connected in parallel with the electrometer. Let C_0 be the capacitance of this system: its value can only be found experimentally. The key K_0 enables the second and insulated pair of quadrants to be earthed when necessary. This key consists of an insulated piece of wire bridging two holes drilled in blocks of paraffin wax and containing a small quantity of calcium chloride—Fig. 38·4 (*b*). After a short exposure to the atmosphere the chloride becomes moist and conducting. The insulated quadrants may be connected by closing the key K_1 to one pole of a battery B, the other pole being earthed. C is a standard parallel plate air condenser: one plate is earthed while the other may be connected to the insulated quadrants of the electrometer by closing the key K_2. Let K_2 be open and K_1 closed, the deflexion of the electrometer needle being θ_1. If Q is the charge on the insulated quadrants and the connexions to them, C_0, their capacity, is given by

$$Q = C_0V = \alpha C_0\theta_1.$$

where V is the potential difference applied and α a constant. When K_1 is opened the deflexion is unaltered if the electrometer is in working order, but on closing the key K_2 the deflexion is reduced to θ_2, the charge Q being shared between the capacities C_0 and C. Then

$$Q = \alpha(C_0 + C)\theta_2.$$

$$\therefore\ (C_0 + C)\theta_2 = C_0\theta_1,$$

or

$$C_0 = C\frac{\theta_2}{\theta_1 - \theta_2}.$$

To Determine the Capacitance of a Small Condenser, a Standard Air Condenser being Available.—Suppose that the capacitance of the electrometer has been determined as above—we shall find it desirable to make this as small as possible and hence the plates of C′ should not be close together. Let C_1, Fig. 38·5, be one condenser of known capacitance and C_2 a second

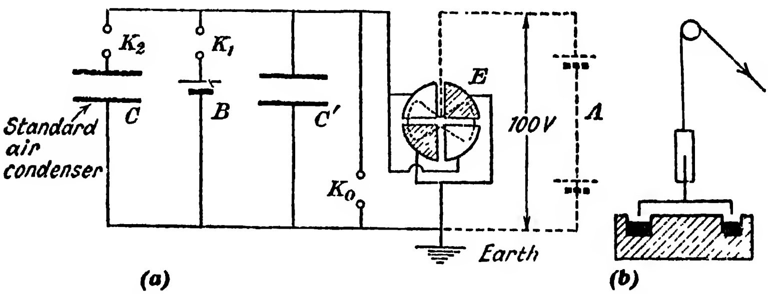


FIG. 38·4.—Capacitance of a Quadrant Electrometer and its Connexions.

condenser arranged so that it may be connected in parallel with C_1. Keys K_0, K_1, and K_2 are arranged as shown.

Let K_0 and K_1 be closed so that C_1, C and the insulated quadrants are raised to a potential V, viz. that of the cell B, and let θ_1 be the electrometer needle deflexion. Then, with the previous notation,

$$Q = \alpha(C_1 + C_0)\theta_1.$$

Now let K_1 be opened and K_2 closed so that the above charge is shared with C_2. Then

$$Q = \alpha(C_2 + C_1 + C_0)\theta_2.$$

Since both C_1 and C_0 are known, C_2 may be determined: in fact

$$C_2 = (C_1 + C_0)\frac{\theta_1 - \theta_2}{\theta_2}.$$

If C_0 is small compared with C_1,

$$C_2 = C_1.\frac{\theta_1 - \theta_2}{\theta_2}.$$

The Permittivity of Ebonite or Glass.—If C_1 and C_2 are two condensers exactly alike so that their capacitances are the same when air is the dielectric, the permittivity κ of a solid may be found by selecting the size of the solid so that it just completely fills the space between the plates of the condenser C_2. The two condensers may be compared as above when the ratio $\frac{C_2}{C_1}$ gives the permittivity or dielectric constant of the solid, since $C_2 = \kappa C_1$.

If the above adjustment cannot be made, or if C_0 cannot be neglected, the following method may be used.

Boltzman's Method for Determining the Dielectric Constants of Solids.—This method resembles an earlier one due to Faraday. A parallel plate condenser is used and the substance under test is in the form of a parallel slab. Let C_1, Fig. 38·5 (*a*), be a fixed air condenser. Its lower plate is earthed while its upper plate is connected to the insulated quadrants of the electrometer E. By means of a key K_0 this part of the system may be earthed when necessary: when K_0 is open and K_1 closed the system is charged by means of the cell B, arranged as shown. C_2 is the experimental condenser arranged as

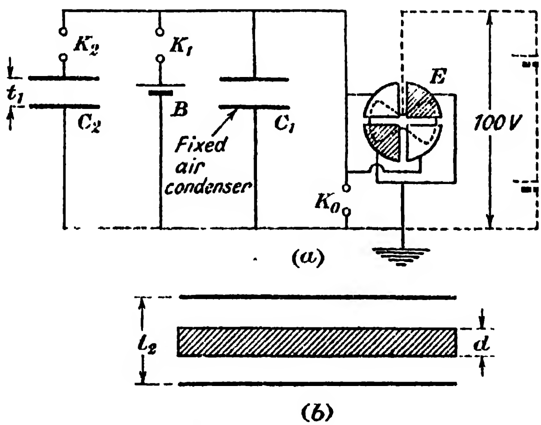


Fig. 38·5.—Boltzman's Method of Determining the Permittivity of a solid.

shown: initially it is uncharged. It may be connected in parallel with C_1 and the electrometer by closing the key K_2. Let θ_1 be the steady deflexion of the electrometer needle, the charging wire having been removed. By closing K_2 the charge on C_1 and the electrometer is shared with C_2—let the steady deflexion of the electrometer needle be θ_2 under these conditions. Let Q be the value of the charge on the condenser C_1 and the insulated plates of the electrometer. If V is the potential difference across the cell, then

$$Q = (C_0)V = \alpha C_0 \theta_1,$$

where C_0 is the capacitance of the electrometer and its connexions.

When the charge is shared with C_2, we have

$$Q = \alpha(C_0 + C_2)\theta_2.$$

The slab of material under test is then placed in C_2, as indicated in Fig. 38·6 (*b*). The distance apart of the condenser plates is then altered to t_2 until the deflexion of the electrometer needle is again θ_2 when the charge on C_1 and the electrometer is shared with the compound condenser. Its capacity must then be C_2. The following analysis shows that the dielectric constant of the material of the slab may be calculated without any knowledge of the values of C_0 or C_2.

Let t_1 be the distance apart of the plates of the condenser C_2 when the dielectric is air: let t_2 be this distance when a slab of uniform thickness d is introduced and the capacitance of this compound condenser restored to its initial value. In this latter instance let E be the electric intensity in the air; then $\frac{E}{\kappa}$ is the electric intensity in the dielectric. Hence V, the potential difference across the condenser, is given by

$$V = E(t_2 - d) + \frac{E}{\kappa}d = E\left[t_2 - d\left(1 - \frac{1}{\kappa}\right)\right].$$

Now $E = 4\pi\sigma$, where σ is the surface density of the charge on the positive plate of the condenser. The capacitance per unit area of the compound condenser is therefore

$$\frac{\sigma}{4\pi\sigma\left[t_2 - d\left(1 - \frac{1}{\kappa}\right)\right]} = \frac{1}{4\pi\left[t_2 - d\left(1 - \frac{1}{\kappa}\right)\right]}.$$

If A is the area of each plate of the condenser, its capacitance is, so far as end-effects may be neglected,

$$\frac{A}{4\pi\left[t_2 - d\left(1 - \frac{1}{\kappa}\right)\right]}.$$

But this is equal to C_2, viz. $\frac{A}{4\pi t_1}$, if end effects are neglected.

$$\therefore\ t_2 - d\left(1 - \frac{1}{\kappa}\right) = t_1$$

or

$$\kappa = \frac{d}{d - (t_2 - t_1)}$$

The Measurement of a Small Electric Current.—If two plates of an air condenser are maintained with a potential difference across them and the air between them is exposed to the action of X-rays or other ionizing agent a small current flows between the plates. This current is too small to be detected by a galvanometer. One method of measuring such a current is as follows:—A very high resistance R, Fig. 38·6, is joined in series with a battery of from 30 to 300 volts or more and a condenser C. When the air between the plates of the condenser is exposed to X-rays a current flows in this circuit, thereby creating a potential difference between the ends of R. These are connected to the opposite

pairs of quadrants of an electrometer and the steady deflexion of the needle is recorded. Let this be θ_1. The electrometer is calibrated by placing a Daniell cell [1·08 volts] across its diagonally

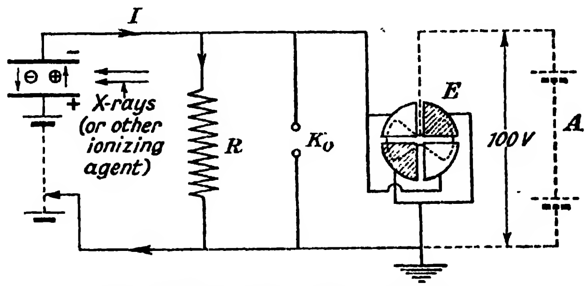


FIG. 38·6.—Measurement of an Ionization Current.

opposite quadrants. Let the deflexion be θ_2. Then the P.D. corresponding to θ_1 is $1{\cdot}08 \times \left(\frac{\theta_1}{\theta_2}\right)$. The current in the circuit is therefore $\frac{1{\cdot}08}{R} \times \frac{\theta_1}{\theta_2}$ amperes if R is measured in ohms. For the experiment to be successful R must be of the order 10^{10} ohms.

The Electrophorus.—This simple piece of apparatus, originally devised by VOLTA, enables an almost infinite number of charges to be obtained from a single initial charge. It consists of a brass plate attached to the under surface of a disc of ebonite. This plate is termed the ***sole***. A second brass plate, to which there is attached an insulating handle, rests upon the upper surface of the ebonite; usually this plate is smaller than the ebonite. A negative charge is given to the ebonite by rubbing it with fur when the lines of force are somewhat as shown in Fig. 38·7 (*a*). [In the ebonite they are lines of electric displacement, cf. p. 712]. The metal disc is then brought near to the charged ebonite surface: actually it is allowed to touch the surface, but owing to the irregular nature of the surfaces, contact is made between them only at a few points. The negative charge on the ebonite charges the metal disc by induction: the field is shown in Fig. 38·7 (*b*) when, for convenience, the distance between the insulated brass disc and the upper surface of the ebonite is greatly exaggerated. The upper disc is then earthed, so that the induced negative charge escapes to earth—cf. Fig. 38·7 (*c*). When the plate is raised as in Fig. 38·7 (*d*) it retains its positive charge, which can be transferred to a suitably arranged condenser. The process is then repeated. It is sometimes necessary to renew the charge on the ebonite, since the initial charge is slowly dissipated especially if the humidity of the

air is high. The labour of touching the second brass plate with the finger at each repetition of the above process may be avoided by having a brass pin passing from the sole to the upper surface of the ebonite so that it touches the plate each time it is placed in position on the ebonite. This permits the negative charge (the electrons) to escape to earth and the state of affairs is as shown in Fig. 38·8 (*c*).

Since the original charge on the ebonite is not diminished by the above process it is of interest to inquire the source of energy. It is found that more work is required to lift the plate when it is charged, for it is then necessary to overcome the force of attraction between the charge on the ebonite and that on the plate which is

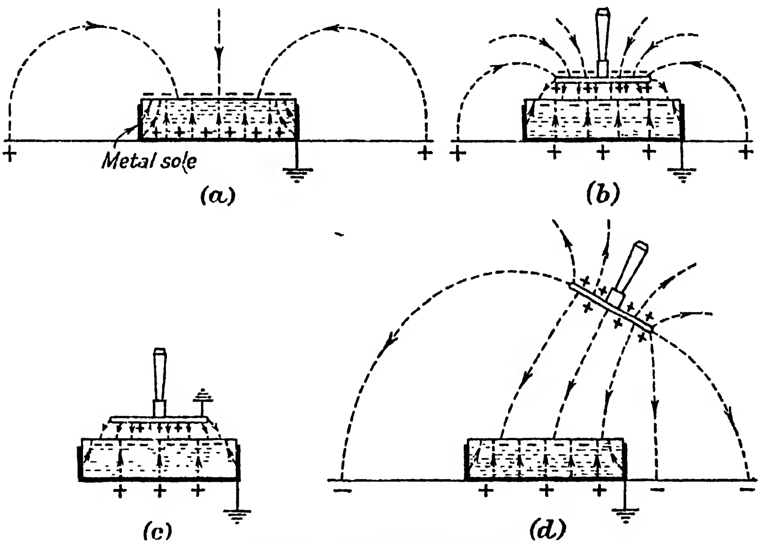


FIG. 38·7.—The Electrophorus.

raised. Hence the source of energy is the extra mechanical work done when the plate is charged.

The part played by the metal sole is somewhat as follows. The negative charge on the ebonite induces positive and negative electricity on the sole, but the latter escapes to earth if the instrument lies on a table. The positive electricity on the sole causes the negative electricity on the ebonite to penetrate slightly into the interior of the ebonite and thus diminish the rate of loss of the charge on the ebonite.

The Wimshurst machine described below is an agency whereby the turning of a handle causes the various stages of the process just described to be repeated cyclically.

The Wimshurst Machine.—The type of influence machine most frequently used consists of two glass plates which have been varnished with shellac. Tinfoil strips are placed radially on the outer sides of these two plates; these plates are capable of being rotated in opposite directions about a horizontal axis. The manner in which such a machine is used is best explained by means of Fig. 38·8, in which the plates are replaced by cylinders. Suppose that the rotation of the cylinders is in the direction of the arrows; further, let us suppose that the tinfoil carrier A has acquired a small positive charge. When A is opposite B, which is connected to D by means of copper wire brushes supported at the ends of a brass rod, then a negative charge is induced on B whilst D acquires a positive charge, since B and D really form one conductor. These charges are separated when the contact between the brushes and the discs on which the charges have been developed is broken.

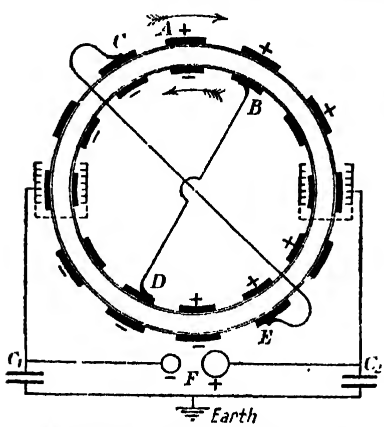


FIG. 38·8.—The Principle of a Wimshurst Machine.

The negative charge on B, moving towards the left, induces a positive charge on C, and a negative one on E. Thus all the strips on the upper half of the outer cylinder acquire positive charges, as do also the strips on the lower half of the inner cylinder. These positive charges pass the collecting combs on the right-hand side of the diagram. These combs are sharp metallic points connected to a knob F, the potential of which is raised as the charge which the combs collect increases. Similarly the smaller knob of the machine acquires a negative charge, so that when the potential difference between the two knobs is sufficiently great a spark passes between them.

The amount of charge collected in the process is small, although the difference in potential between the knobs may rise to 200 e.s.u. or 60,000 V. To increase the charge, the capacitance of the system must be increased. For this purpose two Leyden jars have their inner coatings connected to the discharge knobs while their outer coatings are earthed. These condensers are represented by C_1 and C_2 in Fig. 38·8.

The Condensing Electroscope.—The ordinary gold-leaf electroscope is only suitable for the detection of high voltages. If its disc is

connected to the one electrode of a battery, the other being earthed, then no deflection of the leaves is observed—the applied potential difference is too small. The so-called 'condenser effect,' viz. the raising of the potential difference between the plates of a charged condenser when the distance between them is made greater, may be used to increase the sensitivity of the electroscope. In this instance, the condenser consists of two metal plates A and B, Fig. 38·9 (*a*), each about 20 cm. in diameter and insulated from one another by a sheet of recently dried paper D. A is connected to earth while B is in metallic connexion with the disc E of a gold leaf electroscope, whose outer case is earthed as usual. X is a

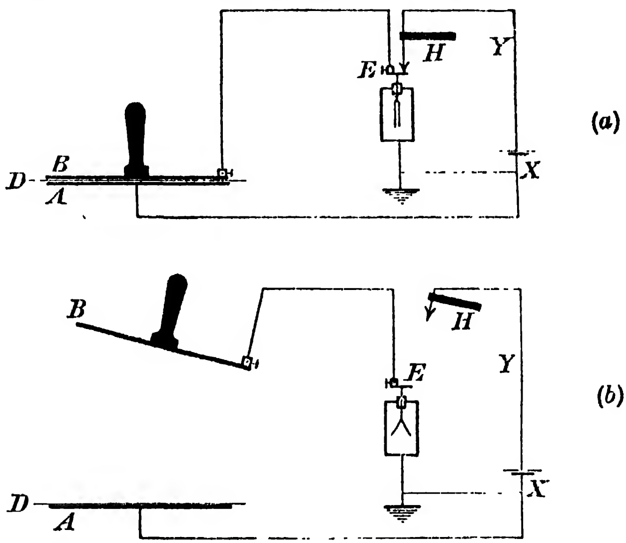


FIG. 38·9.—An Experiment with a Condensing Electroscope to show that there is a Potential Difference between the Terminals of a Voltaic Cell.

battery, one of whose terminals is connected to earth, the other being joined to a wire, Y, to which is attached a stick of sealing wax H; this serves as an insulating handle. When Y touches E the difference of potential between the plates of the condenser is equal to that between the poles of the cell, but the leaves of the electroscope do not diverge. The contact between Y and E is then broken and the plate B raised by means of an insulating handle attached to it—cf. Fig. 38·9 (*b*). The leaves of the electroscope diverge through a considerable angle showing that there is now a large difference in potential between them and the case of the instrument; this is because the capacitance of the condenser has been diminished some hundredfold while the charge on each of its plates remains (except for induction effects) constant.

A blank experiment should always be performed to test whether or not the condenser is charged initially. If it is, it may be discharged by allowing a bunsen flame to pass rapidly over the paper, or by exposing it to X-rays.

EXAMPLES XXXVIII

1.—Describe a quadrant electrometer. For what measurement is it specially suited ?

2.—How may the dielectric constant of ebonite be determined ?

3.—A small current flows through a resistance of 10^{10} ohm., the ends of which are connected to opposite quadrants of an electrometer. The deflexion is 120 scale divisions. When a Daniell cell [e.m.f. 1·08 volt.] is connected across these quadrants, the deflexion is 80 scale divisions. What is the magnitude of the current ?

4.—Describe and give the theory of a 'trap-door' electrometer. Calculate a value for the measured pull on an attracted disc of radius 5 cm., when the insulated plate is 2 mm. away from it and at a potential of 600 volts.

5.—What is meant by electrostatic induction ? Describe the electrophorus and explain how it acts. What is the source of the electrical energy which may be given to a Leyden jar by means of this instrument ?

CHAPTER XXXIX

THE PROPERTIES OF A MAGNET

In many parts of the world there is found a certain oxide of iron, called ***magnetite*** or ***lodestone*** [ἡ λίθος Μαγνῆτις], which has the property of attracting iron filings. The ore is said to possess ***magnetism.*** The name is familiar to all, and yet nobody knows what magnetism really is—the term, like so many others, being really a confession of our ignorance with regard to things which are fundamental. A piece of lodestone, Fe_3O_4, is a ***natural magnet***; the piece of iron which it attracts becomes a magnet too and is called an ***artificial magnet,*** since it now also possesses this remarkable property called magnetism. In these days lodestone is never used for experimental purposes, since artificial magnets can be made which are very much more powerful; but the two types of magnets have identical properties, although the degree to which this property of magnetism is possessed is very different.

Some Preliminary Definitions.—When an artificial bar magnet is dipped into iron filings and withdrawn, it is found that the filings adhere most strongly near the ends of the bar; these regions in which the effects of magnetism are greatest are called the ***poles*** of the magnet. The longer the bar in comparison with its thickness, the more nearly do the poles approach the ends of the magnet. When a ***ball-ended*** magnet, consisting of a steel rod on the ends of which steel balls have been screwed, is magnetized, it acts like a simple magnet with poles at the centre of the balls.

If a steel knitting-needle after being magnetized by stroking it, ***always in the same direction,*** with the pole of a bar magnet, is suspended at its centre by a silk thread, then, when a bar magnet is brought near to the needle, the latter moves. If the end of the bar magnet which was used in the process is brought near to the end of the needle which it finally left, the two are attracted together; placed at the other end of the needle the two are repelled. Evidently the poles of a magnet possess dissimilar properties, i.e. there are two types of magnetism; it is found by experiment that ***similar poles repel one another, while dissimilar poles attract one another.*** Since, however, a magnet will attract a piece of

'unmagnetized' iron, it follows that repulsion is the only sure test for magnetism. The reason for this is given later [cf. p. 741].

When the suspended needle is displaced from its position of rest it continues to execute oscillations for some time, but when these have died down the needle points in its original direction. It is natural to assume that there are some external forces attracting the ends of the magnetized needle. These forces are due to the earth's magnetism, for the earth itself behaves as if it were a large magnet. It is an experimental fact that a suspended magnet points in a direction which is not far removed from that of the geographical north and south. This fact was appreciated by DR. GILBERT, a physician to Queen Elizabeth. The end of the needle which points towards the north is called the ***north-seeking pole*** or ***north-pole*** of the magnet, the opposite end is the ***south-seeking pole;*** the kind of magnetism which is present at one pole of a magnet is referred to as the ***north-seeking magnetism*** [or ***positive magnetism***], while the other is the ***south-seeking*** [or ***negative***] ***magnetism.***

The position has now been reached when the results of the above experiment can be stated more explicitly. If the north pole of a magnet is caused to pass along a needle, the end of the needle which is last in contact with the magnet acquires south-seeking magnetism; the other end is magnetized positively.

If a piece of clock-spring is magnetized by means of another magnet it too becomes a magnet, and has two poles near to its extremities. Suppose now that such a piece of spring is stroked with the positive pole of a magnet starting from the middle and going, in turn, to each extremity. Both the ends are magnetized negatively, whilst there is a positive pole near to the centre of the spring. Such an arrangement may be regarded as a double magnet with the positive pole of one in contact with that of the other—two adjacent like poles in a magnet constitute what is generally termed a ***consequent*** pole.

If a magnet is brought near to a small pivoted magnet—usually termed a compass needle—the small needle is deflected from its position of rest. This must be due to external forces, and it is natural to suppose that the large magnet is responsible for them. The region in which the influence of a magnet can be detected is called a ***magnetic field;*** obviously, the more sensitive the detecting instrument the larger the field which can be observed. Hence, mathematically, it is correct to regard the field of a magnet as infinitely large; practically, it is confined to a small region near the magnet, for its influence cannot be detected beyond these confines.

Induced Magnetism.—Let NS, Fig. 39·1, be a bar magnet supported vertically in a clamp. If a small piece of steel is placed near to S it is attracted by the magnet and if allowed to come sufficiently near remains clinging to the magnet when its support is withdrawn. A second piece may also be supported in a similar way if placed below the first small piece providing the magnet is strong enough. On detaching the first small piece of steel carefully the second will remain in contact with it. If the experiment is repeated without allowing the steel to come into contact with the magnet the steel will again become magnetized, only to a less extent.

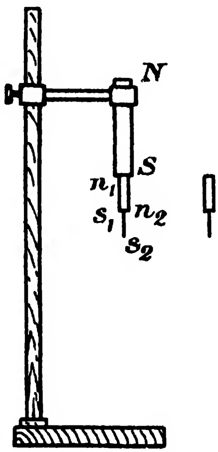


FIG. 39·1.—Induced Magnetism.

The piece of steel magnetized in the above manner is said to have been magnetized by induction and the magnetism in it is referred to as ***induced magnetism,*** although there is no fundamental difference between it and that possessed by the larger magnet. Experiment shows that the nearer ends of the exciting and the excited magnet are opposite in polarity. This fact is readily verified by using the test of magnetic repulsion.

Permanent and Temporary Magnetism.—If the experiment described above is repeated with pieces of soft iron instead of steel, the pieces will still be attracted by the magnet, but when they are removed from the influence of the exciting magnet they will no longer remain together. The reason for this is that soft iron loses the greater part of its induced magnetism when removed from the presence of the inducing magnet. The magnetism it had whilst in contact with the magnet is termed ***temporary,*** while the magnetism it retained when withdrawn from the magnet is called ***permanent.*** In steel the temporary magnetism is practically equal to its permanent magnetism, but with iron the two are widely different.

The Demagnetizing Effect of Magnetic Poles.—The phenomenon of induced magnetism explains the demagnetizing effect of a magnet on itself. When the magnet is in the form of a bar the magnetic force in the bar tends to magnetize, by induction, the material at the centre of the bar. The polarity of this magnetism will be such that south-seeking magnetism is towards the north pole of the magnet and north-seeking towards the south pole, so that the distribution of the induced magnetism is exactly opposite to that of the magnet itself. This self-demagnetizing effect is greatly mini-

mized by the use of soft iron 'keepers' placed across the ends of pairs of magnets as in Fig. 39·2. With magnets bent to form an almost closed ring the demagnetizing effect is automatically reduced since the opposite poles are so close together that the field due to them at the centre of the magnets is comparatively small; consequently its inducing action is small.

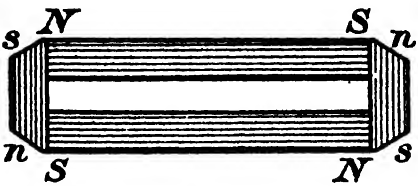


FIG. 39·2.—Bar Magnets with Keepers.

The Making of Magnets.—The inducing action of a magnet on a piece of steel is utilized in constructing small magnets. The three usual methods are known as those of ***single touch, double touch,*** and ***divided touch.*** A description of them will be found in more elementary books than this. These methods are not suitable for

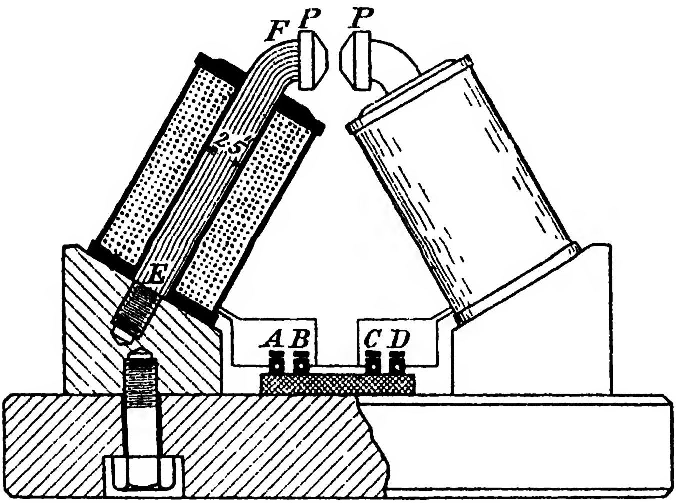


FIG. 39·3—An Electromagnet.

the construction of powerful magnets. To produce these, use is made of the fact that when an electric current is passed through a coil of wire wound round an iron core the iron becomes a very strong magnet, the combination being termed an ***electromagnet.*** A modern form of electromagnet is shown in Fig. 39·3. The core, E, of the magnet consists of special soft magnet steel 2·5 in. in diameter. The two halves of the core are screwed into an iron base and each is surrounded by a coil of copper wire. A, B, C, and D are four terminals, the battery being connected to A and D while a piece

of wire connects B to C. The distance between the poles, PP, may be varied. The field may be so strong that all iron parts must be securely screwed in position before the magnet is excited. When the semi-angle of the pole pieces is about 55° the strength of the field is a maximum for a given current, while if flat poles are used the uniformity of the field is greatest.

To produce a strong permanent magnet the piece of steel is placed symmetrically in a solenoid as in Fig. 39·4. The electrical circuit consists of a key K and a fuse F, this being a piece of wire which melts when the current through it exceeds a certain value, say 10 amperes. The terminals A and B are connected to the mains supplying direct current. On pressing the key K a momentary, but very heavy, current flows. The fuse is blown and the circuit becomes dead, but the current has caused the steel to become highly magnetized.

The Removal of Magnetism.—It is very frequently necessary to remove the magnetism from a magnet. If a body only approximately free from magnetism is required, the magnetism may be destroyed by hammering the specimen or allowing it to fall on the floor, i.e. the specimen must be subjected to mechanical shocks.

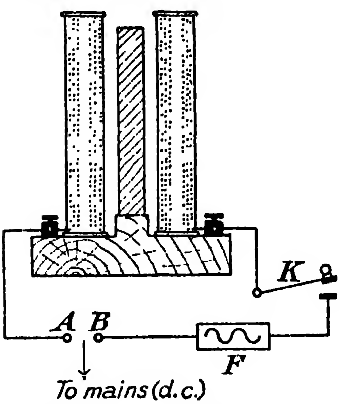


FIG. 39·4.—The Making of a Permanent Magnet.

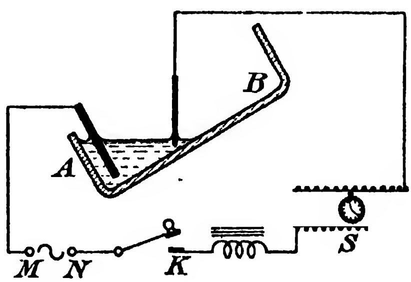


FIG. 39·5.—Arrangement for Demagnetizing the Mainspring of a Watch.

A more effective means is to raise the magnet to a red heat: on cooling, the specimen will be found free from magnetism. Sometimes, however, as for example when the main-spring of a watch has become magnetized, it is not possible to demagnetize the specimen in the above ways. The following arrangement [cf. Fig. 39·5] is always effective. AB is a trough containing a saturated aqueous solution of zinc sulphate [any other conducting solution will serve], the trough being tilted so that the bottom of the trough is only partly

covered. Two electrodes dipping into this solution are connected through a solenoid, S, and a key, K, to a source of alternating current, MN. The watch or other article is placed inside the solenoid. The two electrodes, after being brought near together so that a large current passes in the circuit, are gradually moved farther apart so that ultimately the current is reduced to zero when the specimen will be demagnetized. For the reason for this cf. Chap. L. To be quite certain that the current has been brought continuously to zero it is advisable to splash the solution about in the trough before commencing operations. The moving electrode is finally brought out of the solution by dragging it along the bottom of the trough so that only a very thin film conducts the current.

An alternative method of demagnetizing the main spring of a watch is as follows. An electromagnet with a straight iron core is energized from an a.c. supply: the spring is brought up very close to one end of the magnet, so that it lies in a fairly strong alternating magnetic field, and it is then slowly withdrawn to a region where the magnetic field is inappreciable. During this procedure the spring is subjected to an alternating magnetic field decreasing practically to zero strength so that the spring becomes demagnetized.

Paramagnetic, Diamagnetic, and Ferromagnetic Substances.—In 1845, FARADAY showed that many substances were affected by a magnetic field. Solid specimens were suspended by a long and very fine suspension between the poles of an electromagnet.

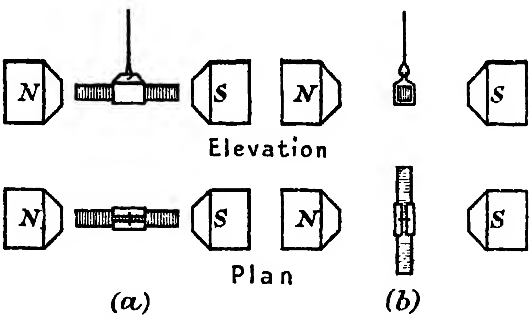


FIG. 39·6.—Paramagnetism and Diamagnetism.

When the current was passed Faraday found that all substances could be divided into two classes. The members of the first class arranged themselves so that their lengths were parallel to the field while the others set in a direction at right angles to the field. These two types are indicated in Fig. 39·6. Faraday called the two classes

paramagnetics and ***diamagnetics*** respectively. When subjecting liquids and gases to this test they were enclosed in narrow glass tubes: the results were, of course, corrected for the magnetic character of the glass.

The first or paramagnetic class comprises substances such as iron, steel, cobalt, nickel, tungsten, aluminium, manganese, and chromium, while bismuth, zinc, copper, lead, and tin are diamagnetics.

Of all the paramagnetic bodies, iron, steel, cobalt, nickel, permalloy, mumetal, and certain alloys known as Heusler's alloys [cf. below], possess the property of becoming very powerful magnets—they are said to be ***ferromagnetic*** substances. In fact, the degree of magnetism possessed by all other substances is so small that it is usual to regard them as non-magnetic, although, strictly speaking, all substances, including gases, are magnetic.

It has also been found that feebly paramagnetic substances behave like diamagnetics when they are placed in a more highly magnetic medium. For example, if a glass tube containing a weak aqueous solution of ferric chloride [Fe_2Cl_6] is placed in a strong magnetic field the tube comes to rest along the lines of force, but if it is supported in a stronger and therefore more highly magnetic solution of the same salt it comes to rest in a direction perpendicular to the field.

Alloys having some Peculiar Magnetic Properties.—In 1892 it was found that although ferro-manganese and ferro-aluminium are only paramagnetic, certain alloys containing about 12 per cent. of iron, the remainder being aluminium and manganese, are ferromagnetic. A year later HEUSLER showed the addition of aluminium, tin, or arsenic, in certain proportions, to an alloy of copper and manganese formed a ternary alloy[1] which was ferromagnetic. The copper-manganese-aluminium alloy is the best known of these so-called ***Heusler's alloys.***

The Heat Treatment of Steel for Use as Magnets.—When steel is heated to a brilliant red heat, and afterwards quenched by plunging it into water, or oil, it becomes very brittle and is known technically as ***glass-hard*** steel. On raising the temperature to a very dull red heat the steel assumes a ***straw tint***: if the heating is continued the tint becomes ***blue.*** Such steel is said to have become ***tempered*** by heat treatment. It is found that steel tempered down to a blue tint retains its magnetism better if used in the construction of magnets having a length more than twenty times their diameter. On the other hand, short magnets have greater retentivity if made from the glass-hard variety of steel.

[1] An alloy having three main constituents.

Cobalt-steel Magnets.—When a powerful magnet is brought near to soft iron filings these become magnetized by influence. If, now, a powerful magnet is brought near to a weak one so that like poles are nearest together, the induced magnetism in the feebly magnetized needle is greater than that originally present, so that attraction ensues. Cobalt-steel magnets are such that the induced magnetism is generally small compared with the permanent magnetism. This enables magnetic repulsion to be demonstrated. An experiment has been described recently in which one magnet is made to 'float' in air like Mohammed's coffin. A cobalt-steel magnet, about 10 cm. long and 0·5 cm. diameter, is placed between two parallel and vertical pieces of glass, on a table. If a second cobalt-steel magnet is placed above the first one so that like poles are together, a considerable force of repulsion is experienced; if the second magnet is released it is seen to float. The glass walls simply serve to prevent the floating magnet from rotating when unlike poles would be brought nearer together and attraction follow.

A Magnetic Balance for Investigating the Inverse Square Law for Magnetism.—This apparatus was originally invented by HIBBERT, but in its simplest form has many disadvantages, the chief one being that the effects of four poles have to be considered. BATEMAN has considerably improved this balance, the principle of which is as follows. AB, Fig. 39·7 (*a*), is a ball-ended magnet

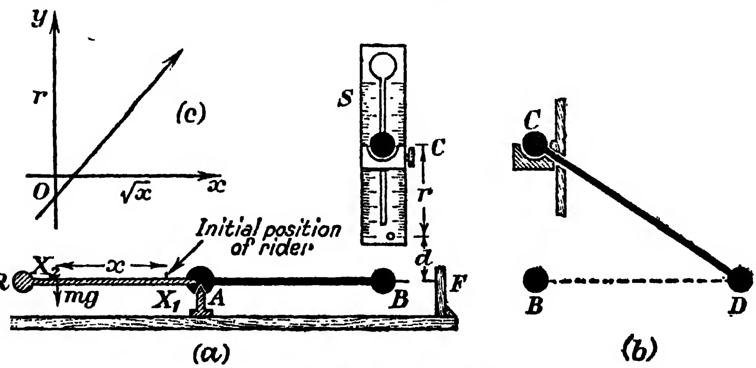


FIG. 39·7.—Magnetic Balance.

pivoted about a horizontal axis through the centre of A. It is counter-balanced by a brass rod R, and a small rider of mass m which may be moved along the rod. To mark the initial position of B a small brass pointer is fixed to it and the position of m is adjusted until the above pointer is opposite a fixed fiducial mark F. CD, Fig. 39·7 (*b*) is another ball-ended magnet placed in a plane perpendicular to the vertical plane through AB, and in such a

position that C is vertically above B while D is in the horizontal plane through the centre of B. Suppose that B and C are like poles of strengths m_1 and m_2. Let the centre of C be r cm. from the zero on a vertical scale S, this zero being d cm. above the centre of B—d is unknown. The force of repulsion between C and B is $\frac{m_1 m_2}{(r+d)^2}$, if the inverse square law is true. C exerts a force on A and D one on A and B, but these forces cannot affect the equilibrium of AB owing to the manner in which the poles are arranged. Thus AB experiences a moment $\frac{m_1 m_2}{(r+d)^2}.\text{AB}$ and B tends to move downwards. Let AB be restored to its zero position by moving m from X_1 to X_2 a distance x along AB. Then, for equilibrium,

$$mgx = \frac{m_1 m_2}{(r+d)^2}.\text{AB},$$

or

$$x(r+d)^2 = \text{constant} = k^2(\text{say}).$$

$$\therefore (r+d) = kx^{-\frac{1}{2}}.$$

If therefore we plot $x^{-\frac{1}{2}}$ against r [cf. Fig. 39·7 (c)], we should obtain a straight line if the inverse square law is valid; the intercept on the y-axis being $-d$.

The above method of investigating the inverse square law for magnetism is of recent date—the apparatus is only used for teaching purposes. Historically, COULOMB used the torsion balance to investigate the above law; about 1833 GAUSS made some very careful measurements in connexion with this law—his work will be described later [cf. p. 766].

Pole-strength.—If two magnets exactly alike were placed at CD, the force of repulsion would be doubled; three magnets and the effect would be trebled. When these results are contemplated one is led to conceive of the idea of a quantity of magnetism, or ***pole strength*** of a magnet. Experiment shows that if two poles of strength m and m' respectively, are separated by a distance r, then F, the force acting on either pole, is proportional to $\frac{mm'}{r^2}$.

This may be written $F = \alpha . \frac{mm'}{r^2}$ where α is a constant, arbitrarily chosen as unity for a vacuum [or air]. Thus, in air, $F = \frac{mm'}{r^2}$.

This equation is really very important, for it is the basis from which the definition of a ***unit pole*** or ***unit pole-strength*** is derived. The unit north-seeking [positive] pole, or unit south-seeking [negative] pole, is ***that pole which, when separated in air by a distance of one centimetre from an equal pole, is repelled by a force of one dyne.***

Also *if a unit pole at a distance of 1 cm. in air from a pole of a magnet experiences a force m dynes due to that pole alone, then the magnet has a pole strength m.*

Magnetic Field Strength or Magnetic Intensity.—The *strength of a magnetic field* or the *magnetic intensity* at a point is numerically equal to the force in dynes which a unit positive pole would experience if placed at that point, it being assumed that the introduction of the unit pole does not alter the configuration of the field. Since the introduction of a unit charge into a field would disturb that field, it is better to define the magnetic field strength, H, by the equation $H = \lim_{\delta m \to 0} \frac{\delta F}{\delta m}$, where δF is the small force experienced by a quantity of magnetism δm, introduced into the field at that point where the magnetic intensity is being considered. Moreover, this equation shows that the dimensions of magnetic intensity are not those of a force, but those of a force divided by a pole strength. The unit of magnetic intensity is the ***oersted,*** and a magnetic field is said to have unit strength when the force acting on a unit positive pole in it is one dyne. To determine the magnetic intensity due to a pole of strength m at a point distant r from it we imagine that a unit positive pole has been placed at the point in question. The force of repulsion between the two poles is $\frac{(m \times 1)}{r^2} = \frac{m}{r^2}$; this gives the magnitude of the intensity; its direction will be along the line joining the two poles and its sense away from the pole m, if this is positive.

Alternatively, if a small positive pole δm at a point distance r from m, experiences a force δF, then $\delta F = \frac{m \cdot \delta m}{r^2}$, i.e. $\frac{\delta F}{\delta m} = \frac{m}{r^2}$. [The sense of the magnetic intensity is given by the sign of δF.] Since $\frac{m}{r^2}$ is the limiting value of $\frac{\delta F}{\delta m}$, it is the magnetic intensity at the point considered.

In general, magnetic field strength or magnetic intensity will be denoted by the symbol H. The symbol H_0 will be used to denote the horizontal component of the earth's magnetic field: its vertical component is H_V.

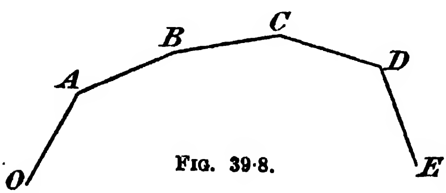
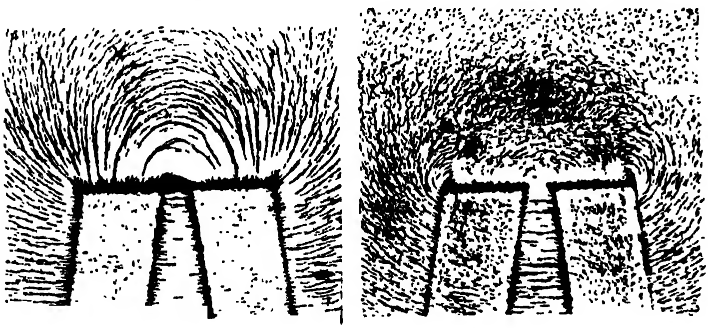


FIG. 39·8.

Magnetic Lines of Force.—Let O, Fig. 39·8, be a point in a magnetic field. Commencing at O let us move a short distance OA in the direction of the magnetic intensity at O. To avoid this somewhat long expression

we frequently say that we have moved in the direction of the field at O. From A let us move another short distance in the direction of the field at A, and so on. In the limiting case when the short distances become infinitely small the broken curve becomes continuous and it has the property that the tangent at any point on it indicates the direction of the field at that point. Such a line is called a ***line of force.*** If a unit pole were placed in a field and released it would move along a line of force provided that sufficient frictional forces were present to prevent it acquiring an appreciable amount of momentum. To plot the lines of magnetic force due to the combined effect of a bar magnet and the earth's field a small compass needle is placed in the field and the positions of its extremities indicated by dots. The needle is then moved to such a position that its S-pole comes to rest over the point previously occupied by its N-pole. Another dot is obtained and the process



(a) (b)

FIG. 39·9.—Lines of Force indicated by Iron Filings.

continued. The curve is obtained by joining successive dots together. Such a method can only be used when the lines of force are not sharply curved, for a compass needle of finite length necessarily lies along the tangent to the line of force at its centre, and in the above process of plotting a field it is tacitly assumed that the tangent coincides with the line of force over a length equal to that of the needle.

A very rapid and interesting way of showing lines of magnetic force consists in laying a piece of sensitized paper on the magnet and sprinkling over it some iron filings, a process which is most readily accomplished by stretching a piece of coarse muslin over the mouth of a bottle containing the filings, and using it as a pepper-box. By gently tapping the paper the filings are caused to arrange themselves along the lines of force. The paper is then exposed to sunlight, the filings are removed, and a permanent record is obtained by fixing the paper in the usual way.

The distribution of the lines of force by means of iron filings is shown in Fig. 39·9 (*a*) and (*b*). In (*a*) the keeper has been removed, while in (*b*) the keeper has been placed near to the poles. The marked absence of the lines of force above the keeper shows that the lines of force prefer to follow the path through the soft iron rather than through the air.[1]

Uniform and Radial Magnetic Fields.—A magnetic field is said to be uniform when the magnitude and direction of the strength of the field are constant at all points in it.

When the direction of the field strength has a common origin the field is radial, e.g. the magnetic field due to a single pole. The term is also applied to magnetic fields in which the direction of the lines of force at all points in it originate on a straight line. In each instance the field has the same numerical value at points equidistant from the point or line origin.

The Effects of Magnetism on Chronometers.—The accuracy of an ordinary watch, having a bimetallic [steel and brass] balance wheel and a steel hair-spring, is greatly affected by magnetism. When placed in a strong magnetic field the steel portions become magnetized and the period is affected since the earth's magnetic field exerts an additional couple on the wheel. Moreover, the hair-spring may be drawn out and touch the wheel. The watch then behaves erratically, and it must be demagnetized.

An elinvar balance wheel [cf. p. 168] is left uncut, and although it may be magnetized, it loses the magnetism on being removed from the field. The magnetic conditions of balance-wheel wheels and hair-springs made (*a*) of elinvar, (*b*) in the usual manner, are indicated in Fig. 39·10 and 39·11.

Verification of the Inverse Square Law.—By constructing a line of force due to a bar-magnet placed in a horizontal position in the earth's field the inverse square law may be verified as follows :—By means of a compass needle draw the lines of force due to the horizontal component of the earth's field alone—they are represented by the parallel lines H in Fig. 39·12. Next place the magnet in any convenient position and construct several lines of force, one of which, viz. NPS, is shown. The point P is selected where the line of force is parallel to the earth's horizontal field. If this condition is complied with, the direction of the total field at P due to the magnet alone must also be parallel to the lines H. But we can determine the direction of the field at any point due to the magnet alone as follows :—Let PN and PS be called r_1 and r_2 respectively. Then

[1] The 'line of force' is purely a mathematical concept, but it is of such use in explaining magnetic phenomena that we wonder it does not correspond to some reality and we talk as if it does.

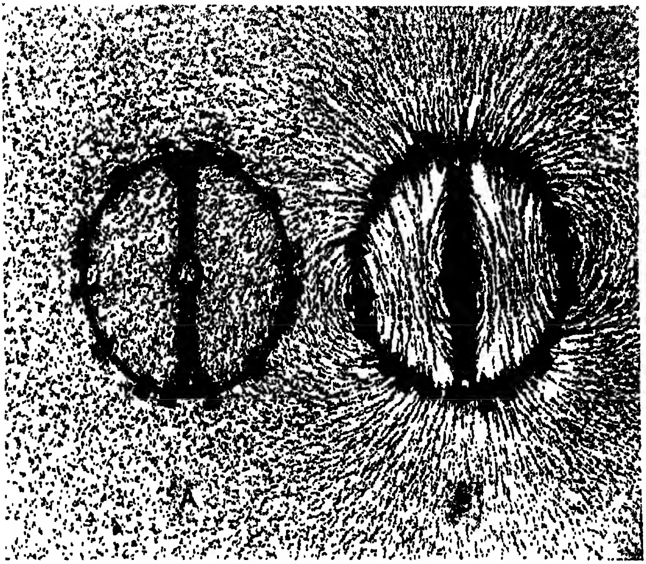


FIG. 39·10.—Magnetic Conditions of Balance Wheels.
A. Elinvar. B. Steel and Brass.

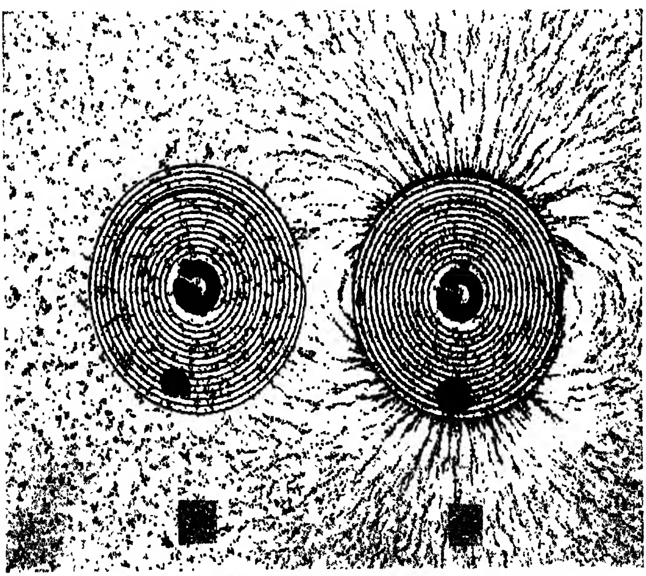


FIG. 39·11.—Magnetic Conditions of Main-Springs.
A. Elinvar. B. Steel.

the total field at P due to NS alone has two components, numerically equal to $\frac{m}{r_1^2}$ along NP, and $\frac{m}{r_2^2}$ along PS. We therefore draw PA and PB proportional to these components and complete the parallelogram PACB. If its diagonal PC, which represents the total intensity at P due to NS alone, is parallel to the lines H the inverse square law will have been verified. For the purposes of this experi-

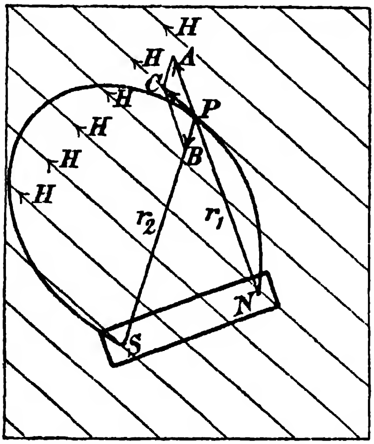


FIG. 39·12.—Verification of Inverse Square Law.

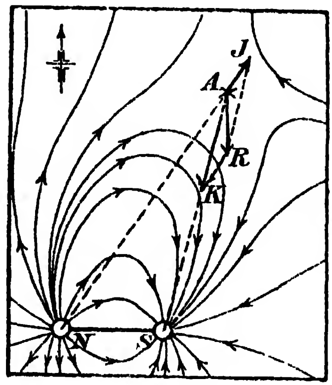


FIG. 39·13.

ment and others—unless ball-ended magnets are employed—it may be assumed that the poles are symmetrically placed and that the distance between them is five-sixths the total length of the magnet. Good results are obtained by using ball-ended magnets at least 15 cm. long, and choosing the point P so that it is at least 15 cm. away from each pole. Moreover, the curvature of the line of force at P should be small [cf. p. 749].

Neutral points.—A ***neutral point*** in a magnetic field is defined as ***a point at which the strength of the magnetic field is zero.*** In actual practice the neutral points which are located experimentally do not, strictly speaking, satisfy the above definition, for they are such that at any one of them the resultant field strength due to the magnet and the earth has no component in any horizontal direction whatsoever, i.e. at such a point the resultant magnetic field, if any, is entirely vertical. Such a point should therefore be termed a ***pseudo-neutral point,*** but where no confusion arises it will, in the sequel, often be designated a neutral point for the sake of brevity and in conformity with general practice.

If a bar-magnet lies with its axis in an E.–W. direction, its

north-seeking pole pointing west, the lines of force are as indicated in Fig. 39·13. At the point A [there is a symmetrically placed other and similar point in the part of the diagram not reproduced] it will be noticed that a compass needle tends to set in any position: A is a neutral point or, better, a pseudo neutral point. In the diagram AJ and AK represent the magnetic field strengths at A due to the N and S poles of the magnet. The resultant magnetic field at A due to the magnet NS is represented by AR, the diagonal through A of the parallelogram whose sides are AJ and AK. Since A is a pseudo neutral point the above resultant field is equal and opposite to H_0, the horizontal component of the earth's magnetic field. It therefore follows that when such a point as A has been located the pole-strength of the magnet, NS, may be calculated if the value of H_0 is known. [In London, $H_0 = 0{\cdot}185$ oersted.] The method is exemplified in the next paragraph.

Experimental Determination of Pole-Strengths and Magnetic Moments.—(*a*) If a magnet has its north-seeking pole N pointing to the north two neutral points are found on the equatorial line of the magnet: they are at equal distances from the centre of the magnet, and O, Fig. 39·14 (*a*), is one of these points at distance r from either pole. The magnetic field strength at

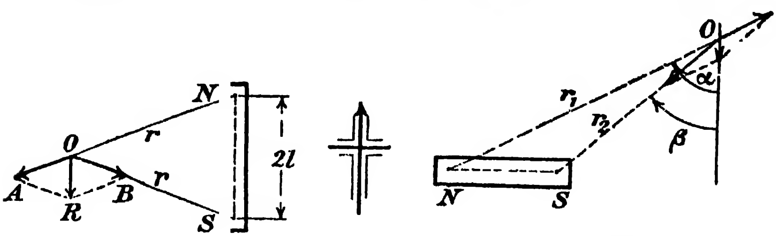


FIG. 39·14.

O due to NS alone has two components each numerically equal to $\frac{m}{r^2}$. If OA and OB represent these in magnitude and direction, OR, the diagonal of the parallelogram OARB represents their resultant. It is equal to $2\cdot\frac{m}{r^2}\cos \widehat{AOR} = 2\cdot\frac{ml}{r^3} = \frac{M}{r^3}$, where $M = 2ml$. This quantity, $2ml$, the pole strength × the magnetic length, is termed the ***magnetic moment*** of the magnet.[1] At a neutral point the above expression is equal and opposite to H_0 so that

$$M = H_0 r^3, \text{ and } m = \frac{H_0 r^3}{2l}.$$

[1] The unit of magnetic moment is the unit-pole.cm., or alternatively, cf. p. 760, the erg.oersted.$^{-1}$

If the direction of the magnet is reversed the neutral points lie on the axis of the magnet and it is left as an exercise to the student to prove that in this instance $M = \frac{1}{2}H_0\frac{(r^2 - l^2)^2}{r}$, where r is the distance of a neutral point from the *centre* of the magnet.

(*b*) When the bar magnet points east and west the two neutral points lie on a line inclined to the axis of the magnet. Let O, Fig. 39·14 (*b*), be one of the neutral points for this position of the magnet. If r_1 and r_2 are the distances NO and SO respectively, while α and β are the angles these vectors make with the direction of H_0, the component in a direction opposite to that of H_0 of the intensity at O due to NS alone, is $\frac{m}{r_2^2}\cos\beta - \frac{m}{r_1^2}\cos\alpha$. Since O is a neutral point the above component is equal and opposite to H_0, so that

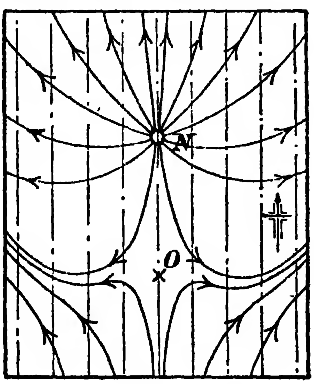
$$H_0 = \frac{m}{r_2^2}\cos\beta - \frac{m}{r_1^2}\cos\alpha,$$

and, in addition,

$$\frac{m}{r_2^2}\sin\beta = \frac{m}{r_1^2}\sin\alpha.$$

Although this last equation cannot be used to determine m, its validity should be verified to show that the position of the neutral point has been located correctly.

If ball-ended magnets are available for these experiments better results will be obtained since the positions of the poles are known more accurately—they are at the centres of the spheres.



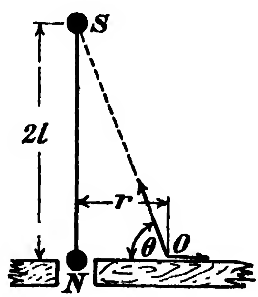


Fig. 39·15.
Magnetic Field due to a Vertical Magnet.

The Magnetic Field Due to a Vertical Magnet.—We may obtain some idea of the configuration of the field in this case without

resort to actual experiment although, if numerical results are to be obtained, the field must be plotted in the usual way. If N, Fig. 39·15 (*a*), is a single north pole the lines of force are straight lines radiating from the pole. If a uniform magnetic field H_0 (the earth's horizontal magnetic field) is superposed on this, the lines of force in the upper half will tend to bend round and become parallel to H_0. The lines of force in the lower half will commence similarly to travel southwards but will gradually bend round as indicated. In the same way the field H_0 will also be disturbed as shown. There will be a neutral point at O. Now in actual practice there will always be present the south pole of the magnet so that the actual arrangement of the lines will be slightly different from that shown. But even so there will still be a neutral point. Let this be distant r from N, Fig. 39·15 (*b*), which is a section through the magnet and the neutral point. The horizontal components of the intensity at O due to the magnet alone are $\frac{m}{r^2}$ and $\frac{m}{(r^2 + 4l^2)} \cos \theta$. Since these act in opposite directions and O is a neutral point we have

$$H_0 = \frac{m}{r^2} - \frac{m}{(r^2 + 4l^2)} \cos \theta.$$

Experimental Verification of the Inverse Square Law for Magnetism.—NS, Fig. 39·16, is a ball-ended magnet placed at random on a table. The lines of force are plotted in the usual way. Let us suppose that O is a neutral point. Then the

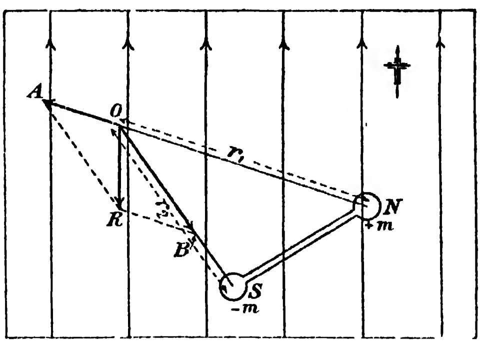


Fig. 39·16.—Experimental Verification of the Inverse Square Law for Magnetism.

field at O due to the magnet alone is parallel to H_0, the horizontal component of the earth's magnetic field, but in the opposite sense.

To verify the *inverse square law for magnetism*, through O, draw OR of any convenient length parallel to H_0 but in the opposite sense, and through R draw RA and RB parallel to SO and NO to

cut NO produced in A and SO in B respectively. Then OA and OB are proportional to the magnetic intensities at O due to the positive and negative poles of the magnet respectively. If the force between two poles in air is inversely proportional to r^n, where n is to be determined, then

$$OA = km/r_1^n$$

and $OB = km/r_2^n$, where k is a constant, m the numerical value of the pole strength in arbitrary units, and r_1 and r_2 are the distances indicated.

Hence

$$\frac{OA}{OB} = \left(\frac{r_2}{r_1}\right)^n,$$

or

$$\log\left(\frac{OA}{OB}\right) = n \log\left(\frac{r_2}{r_1}\right).$$

If therefore the above distances are measured and the value of the expression $\log (OA/OB) \div \log (r_2/r_1)$ is found to be 2, the validity of the inverse square law will have been verified.

Directional Loci in a Magnetic Field and the Locating of Neutral Points.—The methods which have been discussed for locating neutral points in a magnetic field are really very ineffective since, on account of the finite length of the exploring magnet, it is impossible to plot lines of force accurately when their curvature is considerable. OWEN has described the following method. A neutral point is located as the point of intersection of two curves each of which is such that at all points on it the direction of the magnetic field in a horizontal plane is the same. Any two directions may be selected, but it is preferable to have two directions mutually perpendicular, and in cases where the earth's horizontal magnetic field is involved, these two directions should be respectively parallel and perpendicular to this field. Such loci can be determined quite definitely since the errors of plotting are non-cumulative. As the point of intersection of two such ***directional loci,*** as they are termed, is approached the component of the field in the chosen direction diminishes to zero and, beyond it, changes sign. The points of intersection are therefore points at which the horizontal field strength is zero in any direction, i.e. they are neutral points.

The apparatus required to plot directional loci consists of a sheet of squared paper mounted on a drawing-board, and an ordinary small compass needle. The paper is adjusted so that the lines on it run parallel and perpendicular to H_0. When the magnet has been placed in position the compass needle is moved across the paper and some point is soon found where the needle is perpendicular to the magnetic meridian. A sharpened point of a pencil is held firmly just above the centre of the needle, the needle is removed,

the pencil point moved straight down on the paper and a dot made. Other points are found and the locus drawn. The second locus is similarly constructed.

Fig. 39·17 shows two directional loci when a short magnet ('Alcomax steel') lies as indicated in an E.–W. direction. For points on the locus marked (i) the field in the plane of the diagram is parallel to the axis of the magnet; on the locus (ii) the field is perpendicular to the above axis. The loci intersect at X which is a neutral point (really a pseudo-neutral point, cf. p. 752): a similarly situated neutral point would be found if the lower half of the diagram had been completed. Actual measurements give

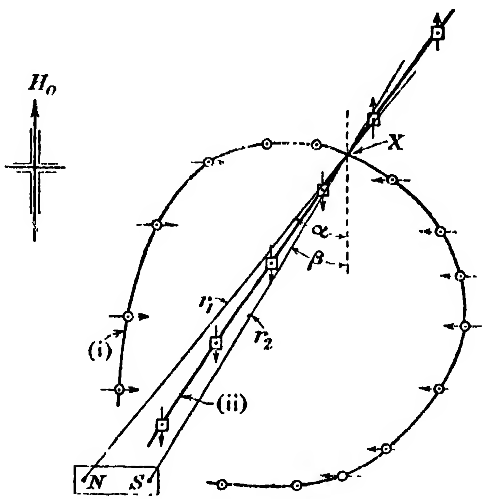


FIG. 39·17.—Directional Loci in a Magnetic Field.

$r_1 = 11{\cdot}2$ cm.; $r_2 = 10{\cdot}2$ cm.; $\alpha = 38{\cdot}75°$; $\beta = 31{\cdot}0°$. Assuming $H_0 = 0{\cdot}18$ oersted, we have

$$0{\cdot}18 + \frac{m}{r_1^2}\cos\alpha - \frac{m}{r_2^2}\cos\beta = 0,$$

whence m, the pole strength of the magnet, is 89 unit-pole.

Also if $2l = 1{\cdot}8$ cm., the magnetic moment of the magnet is given by

$$M = 160 \text{ erg. oersted.}^{-1}$$

It is also interesting to note that

$$\frac{\sin\alpha}{r_1^2} - \frac{\sin\beta}{r_2^2} = (0{\cdot}498 - 0{\cdot}495) \times 10^{-2},$$

$$= (0{\cdot}003) \times 10^{-2},$$

which is negligible compared with either term or the l.h.s. of the

equation so that, with due allowance for experimental error, it may be considered zero, as theory requires.

Neutral points in other fields may similarly be located.

Intensity of Magnetization.—This is defined as the magnetic moment per unit volume of a magnet. If the magnet is uniform in cross-section, and the intensity of magnetization, J, also uniform, then J is equal to the pole strength per unit area of cross-section, for

$$J = \frac{M}{v} = \frac{2ml}{2ls} = \frac{m}{s},$$

where $2l$ is the length of the magnet, v its volume, and s its area of cross-section. It must be noted that $2l$ is now the total length of the magnet, the poles being at the ends of the rod since it is uniformly magnetized.

The unit of intensity of magnetization is the unit-pole.cm. $\div$ cm.^{3}, i.e. unit-pole.cm.$^{-2}$ This is identical with the unit of magnetic induction [cf. p. 791] and likewise may be termed the ***gauss.***

EXAMPLES XXXIX

1.—Calculate the force between magnetic poles of strengths 19 (N) and 27 (S) respectively when separated by a distance of 10·5 cm.

2.—Find the magnetic field strength at a point 6·7 cm. away from a magnetic pole of strength 31·7 unit-pole.

3.—How far away must two like poles of strengths 81 and 54 respectively be placed so that the force between them may be equal to the weight of a 0·50 gm. mass?

4.—A short bar magnet lies in the magnetic meridian. If there is a neutral point 7 cm. from the centre of the magnet calculate the magnetic moment of the magnet assuming the horizontal component of the earth's field to be 0·185 oersted.

5.—Define unit magnetic pole, and explain what is meant by the strength of a magnetic field. Give a short account of the molecular theory of magnetization.

6.—Describe how, in the absence of any external magnetic field, you would proceed to ascertain whether or not one of two identical pieces of iron rod were magnetized.

7.—Describe two distinct tests which can be applied to find out whether a steel bar is magnetized. If it is magnetized explain how it can be demagnetized and, if unmagnetized, describe fully how an electric current can be used to ensure that a particular end of the bar shall have north-seeking polarity.

8.—Describe, with diagrams, the arrangement of a simple circuit suitable for the magnetization of a short steel bar. Show on your diagram the direction of the current and the north pole of the bar.

The bar is removed from the circuit after magnetization and placed on a horizontal table with its north pole pointing S. Show on a diagram the approximate distribution of the lines of magnetic force in the plane of the table. If a neutral point occurs at a distance of 10 cm. from the centre of the bar estimate the magnetic moment of the bar.

[The value of the horizontal component of the earth's magnetic field may be taken as 0·18 oersted.]

CHAPTER XL

MAGNETOMETRY

The Magnetic Moment of a Magnet.—When a small compass needle is placed in a horizontal plane and is free to rotate about a vertical axis passing through its centre it comes to rest in the magnetic meridian if no magnetic materials are present and there are no strong magnetic fields. If the needle is displaced it tends to return to the above position. This motion is caused by two forces acting on the poles of the magnetic needle. If H_0 is the

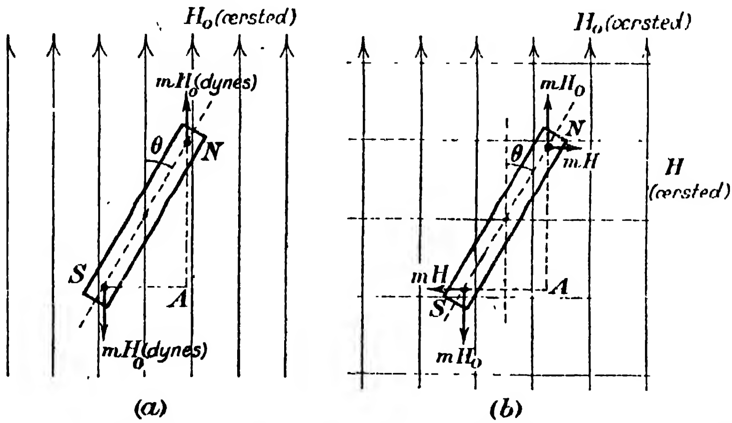


FIG. 40·1.—(a) Couple acting on a Magnet in a Uniform Field when the Axis of the Magnet is not Parallel to the Field. (b) Equilibrium of a Magnet in two Uniform Fields, mutually perpendicular.

horizontal component of the earth's magnetic field and m the pole-strength of the small magnet, the force on each pole is mH_0; but the sense of each force is different since the pole-strengths are really m and $-m$. Let NS, Fig. 40·1 (*a*), be a small magnet displaced from its position of rest through an angle θ, and let $2l$ be the length of the magnet.

The two forces mH_0 constitute a couple, the moment of which is $mH_0 \cdot SA$, where SA is the perpendicular distance between the lines of action of the two forces. But $SA = 2l \sin \theta$, so that the

moment of the restoring couple is $m \cdot 2l \cdot H_0 \cdot \sin\theta = MH_0 \sin\theta$, where $M = 2ml$, the magnetic moment of the magnet [the compass needle]

This equation shows that the ***magnetic moment, M, of a magnet is numerically equal to the couple required to hold it in a position at right angles to a unit magnetic field,*** i.e. to a uniform field of strength 1 oersted. More strictly, ***M is the couple per unit field necessary to hold the magnet at right angles to a small magnetic field.*** The unit is the erg. oersted.$^{-1}$

This is a better definition of magnetic moment than that given on p. 753 and it enables us to define what is meant by the pole strength of a magnet, viz., the magnetic moment of a magnet divided by the distance between its poles.

To Show that a Magnetic Moment is a Vector Quantity.—Let us represent a magnet of magnetic moment M by a straight line OA, Fig. 40·2, along the axis of the magnet, the positive direction of the axis being from the south-seeking to the north-seeking pole of the magnet. Let OP and OQ be the projections of OA on two straight lines mutually perpendicular, the $\widehat{AOP}$ being ϕ. Then OP and OQ will represent magnets whose moments are $M\cos\phi$ and $M\sin\phi$ respectively. Is this procedure legitimate? Let us suppose that H is the magnetic intensity of a uniform field in the plane of the diagram. Let the direction of H make angles θ and ψ with OA and OP respectively. Then the couple on the magnet whose moment is $M\cos\phi$ is $(M\cos\phi).H\sin\psi$, and is anticlockwise. Since $\psi = \theta + \phi$, the above expression becomes

$$MH\cos\phi \,.\, \sin(\theta + \phi).$$

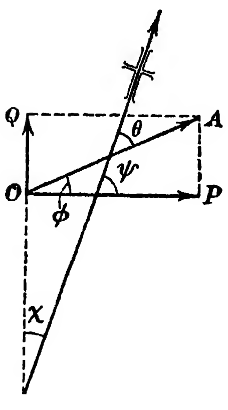


FIG. 40·2.

For the magnet whose moment is $M\sin\phi$, the couple on it in an anticlockwise direction is $-M\sin\phi.H\sin\chi$, where χ is the angle indicated. This is

$$-MH\sin\phi.\sin\left(\frac{\pi}{2} - \theta - \phi\right) = -MH\sin\phi.\cos(\theta + \phi).$$

The sum of these couples is $MH\{\cos\phi.\sin(\theta+\phi) - \sin\phi\cos(\theta+\phi)\}$

$$= MH\sin(\theta + \phi - \phi) = MH\sin\theta,$$

which is the couple on the magnet whose moment is M. Thus it has been found that M is a vector quantity.

The Equilibrium of a Magnet in a Magnetic Field due to the Superposition of Two Mutually Perpendicular Magnetic Fields.—Let us now suppose that the needle is deflected permanently by placing a magnetic field,[1] H, Fig. 40·1 (*b*), at right

[1] This is an abbreviated statement; it means that the magnetic strength of the field, or the magnetic intensity, is H dyne. unit-pole.$^{-1}$ or H oersted.

angles to H_0. The restoring couple on the magnet due to the presence of the second field is $mH \cdot AN$. The equilibrium position of the magnet will be such that restoring couples due to the fields are equal, i.e.,

$$mH \cdot AN = mH_0 \cdot SA.$$

or

$$H = H_0 \tan \theta.$$

Hence, if H_0 is known and θ is measured, H may be deduced. The above relationship is a fundamental one in magnetometry.

[In general, H is only uniform over a small region—hence the magnet used should be short.]

The Magnetic Field Strength or Magnetic Intensity due to a Bar Magnet at a Point on its Axis.—It is required to determine the magnetic intensity [or field] due to a bar magnet at a point on its axis—the axis or axial line of a magnet being defined as the direction of the line joining the two poles together. Let m be the pole-strength, $2l$ the magnetic length of the magnet and r the distance of the point A from the *centre* of the magnet,

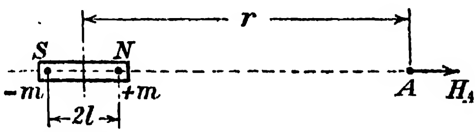


FIG. 40·3.—Magnetic Field Strength at a Point on the Axis of a Bar Magnet.

Fig. 40·3. Suppose that a small pole δm is placed at A. Then the force on this pole due to $+ m$ is $\frac{m \times \delta m}{(r - l)^2}$, since the distance of separation of the two poles is NA or $(r - l)$. This force is considered positive when it acts in the direction of r increasing. The force due to $- m$ is similarly $- \frac{m . \delta m}{(r + l)^2}$. The two forces on the small pole at A act along the direction of r so that their resultant δF [say] is given by

$$\therefore \delta F = \left[\frac{m}{(r - l)^2} - \frac{m}{(r + l)^2}\right]\delta m = \frac{4ml . r . \delta m}{[\,(r - l)\,(r + l)\,]^2}$$

$$= \frac{2Mr . \delta m}{(r^2 - l^2)^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where M is the magnetic moment of the bar magnet.

The field strength at A is the force per unit pole on δm, viz. $\delta F \div \delta m = H_A$.

$$\therefore H_A = \frac{2Mr}{(r^2 - l^2)^2}.$$

If l is small compared with r, i.e. $\frac{l}{r} \rightarrow 0$, the expression for the field strength becomes

$$\frac{2Mr}{r^4\left[1 - \frac{l^2}{r^2}\right]^2} = \frac{2M}{r^3}.$$

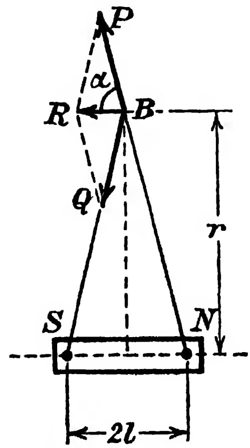


FIG. 40·4.—Magnetic Field Strength due to a Bar Magnet at a Point on its Equatorial Line.

The Field Strength or Magnetic Intensity due to a Bar Magnet at a Point on the Straight Line[1] bisecting its Length at Right Angles.—Let B, Fig. 40·4, be the chosen point at distance r away from the centre of the magnet. The magnetic field strength at B is the resultant of two components (*i*) $\frac{m}{(NB)^2}$ along NB and (*ii*) $\frac{m}{(SB)^2}$ along BS. If these are represented by the vectors BP, BQ, drawn along NB produced and BS respectively, the resultant field strength will be represented by BR the diagonal through B of the parallelogram BPRQ. The magnitude of this resultant is given by

$$H_B = 2BP \cos \alpha = \frac{2m}{BN^2} \cdot \frac{l}{BN} = \frac{M}{(r^2 + l^2)^{\frac{3}{2}}}, \quad . \quad . \quad . \quad (2)$$

since $BN = (r^2 + l^2)^{\frac{1}{2}}$.

When l is small, $\frac{l}{r} \rightarrow 0$, and

$$H_B = \frac{M}{r^3\left[1 + \frac{l^2}{r^2}\right]^{\frac{3}{2}}} = \frac{M}{r^3}$$

Exercise.—Consider a small pole δm at B, Fig. 40·4. Write down expressions for the forces on δm due to the magnet and hence obtain an expression for the field strength.

The Deflexion Magnetometer.—This consists essentially of a small magnetic needle pivoted or suspended by a silk thread, so that it is capable of moving in a horizontal plane. A light aluminium pointer is attached at right angles to the needle, and this is used to determine the angle through which the magnetometer needle moves. The end of the aluminium pointer moves over a circular scale, graduated in degrees. In order to assist the making of accurate observations a mirror is placed underneath the needle, the eye being placed in such a position that the needle and its image are

[1] Sometimes termed the equatorial line of the magnet.

in the plane containing the eye. In this way parallax errors are avoided—see Fig. 40·5 (*a*). This diagram shows that unless the eye is at E_1 directly over the end of the pointer and its image, a considerable error may be made in reading the position of the pointer. [For convenience the scale is shown by vertical lines—actually they are horizontal.] The whole is enclosed in a box furnished with a glass lid protecting the needle from currents of air, etc. Two scales in cm., etc., are fixed, one at right angles to the length of the magnet in its zero position, and the other parallel to it, the centre of the needle being directly over the point of intersection of the axes of the two scales. In other words, these scales point to the (magnetic) east and west, and to the (magnetic) north and south respectively, so that the position of a magnet which is used to deflect the magnetometer needle may be observed [cf. Fig. 40·5 (*b*)].

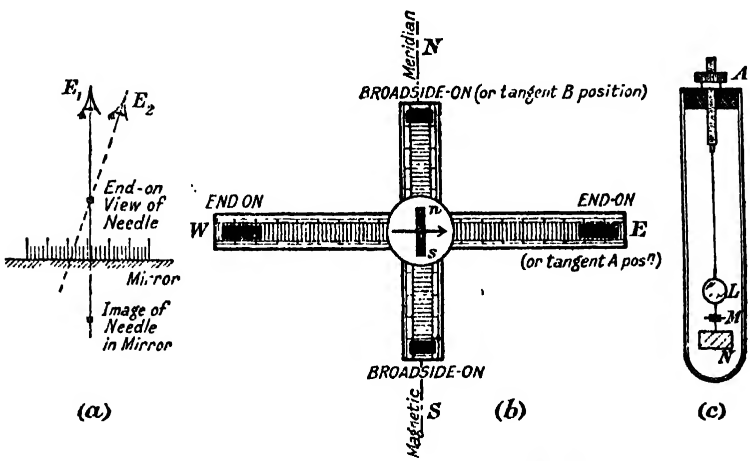


FIG. 40·5.—Deflexion Magnetometers.

A more sensitive type of deflexion magnetometer is indicated in Fig. 40·5 (*c*). It is contained in a wide glass tube about 30 cm. long to protect the actual working part of the instrument from air currents. A No. 2 B.A. brass rod A, fitted through an ebonite disc inserted in the top of the glass tube, supports a fine quartz or silk thread carrying a small concave mirror, L, rigidly attached to a magnet, M, consisting of three short steel rods, and a light aluminium or paper vane, N, the purpose of which is to increase the damping by augmenting the air resistance and thus bring the magnet to rest more quickly after it has been displaced. The deflexions are shown by means of a spot of light reflected from the mirror. The advan-

tages gained by the use of quartz threads are that the restoring couple on the magnet is less than with other forms of suspension, and the elastic properties of quartz are such that after the quartz has been twisted it recovers its former shape completely, a constant zero position thereby being obtained. Very frequently the magnets are mounted at the back of the mirror.

In using a deflexion magnetometer it is customary to arrange that the fields H and H_0 shall be mutually perpendicular, so that the relation $H = H_0 \tan \theta$ is at once applicable.

The condition that the two fields should be perpendicular to one another is not essential although it is one of great convenience. For suppose that the field H makes an angle α with the E.–W. direction. Then the condition for the equilibrium of the magnet is that

$$mH_0 . 2l \sin \theta = mH . 2l \cos (\theta + \alpha)$$

or

$$\frac{H}{H_0} = \frac{\sin \theta}{\cos(\theta + \alpha)},$$

where θ is the deflexion of the needle.

Verification of the Inverse Square Law.—Let AB, Fig. 40·6 (*a*), be a ball-ended magnet having one of its poles directly above the centre of a magnetometer needle. The effect of this pole on the needle will be zero since each pole of the needle is affected in an equal but opposite way by it; moreover, these forces act in a vertical plane, and the needle is only free to swing in a horizontal plane. Hence any deflexion of the magnetometer needle will be due to the pole A. Let this be of strength m and at a distance r away. Then the field at C due to this is $\frac{m}{r^2} = H$ [say]. If θ is the angle of deflexion, m_1, the pole strength of the magnetometer needle, and H_0 the intensity of the earth's horizontal field, we have, from Fig. 40·6 (*b*), $H = H_0 \tan \theta$, or $r^2 \tan \theta = \text{constant}$. If, therefore, when log tan θ (ordinate) is plotted against log r, a straight line whose slope is -2 is obtained, the inverse square law will have been established experimentally.

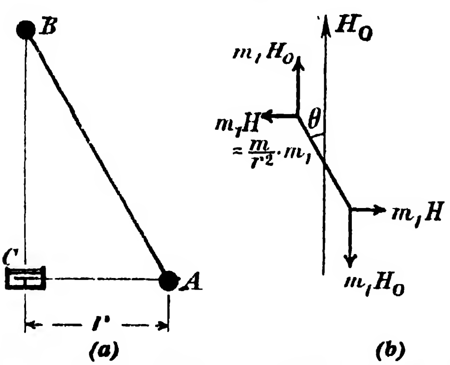


FIG. 40·6.—Verification of Inverse Square Law.

The Tangent A and Tangent B Positions of Gauss.—In Fig. 40·7 let a small compass needle be placed at a point on the axis

of a magnet NS. If H_A is the intensity of the field due to the bar magnet, then H_A is $\frac{2Mr}{(r^2 - l^2)^2}$, and if r is large H_A may be assumed to be uniform over the region occupied by the small compass or magnetometer needle. If H_0 is the value of the horizontal component of the earth's magnetic field, then $H_A = H_0 \tan \theta_A$.

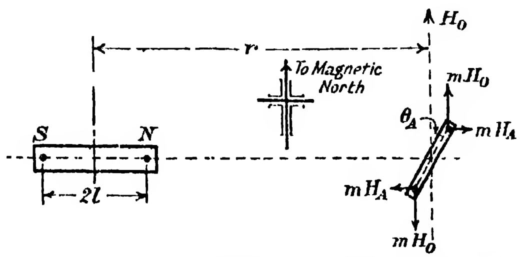


FIG. 40·7.—The Tangent A Position of Gauss.

Substituting the known value of H_A, the equation becomes

$$\frac{2Mr}{(r^2 - l^2)^2} = H_0 \,.\, \tan \theta_A$$

or

$$\frac{M}{H_0} = \frac{(r^2 - l^2)^2}{2r} \,.\, \tan \theta_A . \qquad \text{(i)}$$

If $\frac{l}{r} \rightarrow O$, the above equation becomes

$$\frac{M}{H_0} = \frac{r^3}{2}\left(1 - \frac{l^2}{r^2}\right)^2 \tan \theta_A$$

$$= \tfrac{1}{2} r^3 \tan \theta_A + \text{terms which are negligible}$$

$$= \tfrac{1}{2} r^3 \tan \theta_A \qquad \text{(ii)}$$

Similarly, if the magnet and needle are placed as in Fig. 40·8, then, if θ_B is the corresponding deflexion,

$$H_B = H_0 \tan \theta_B$$

whence, by substitution, and rearrangement of the terms,

$$\frac{M}{H_0} = (r^2 + l^2)^{\frac{3}{2}} \tan \theta_B \qquad \text{(iii)}$$

This reduces to

$$\frac{M}{H_0} = r^3 \tan \theta_B, \text{ when } \frac{l}{r} \rightarrow O \qquad \text{(iv)}$$

These two arrangements of the magnet and needle are called the tangent A and tangent B positions of GAUSS respectively GAUSS was a German mathematician of the early nineteenth century, and the formulation of the above equations was origin-

ally due to his work. These positions are sometimes referred to as the ***end-on*** and ***broadside-on*** positions respectively [cf. Fig. 40·5]. It should be noted that in each position the axis of the deflecting magnet is at right angles to the earth's horizontal magnetic field.

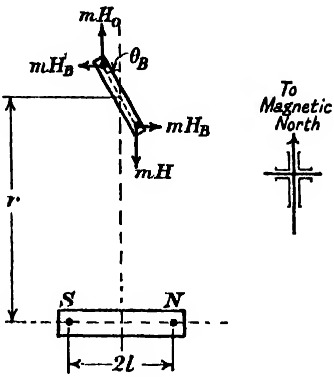


FIG. 40·8.—The Tangent B Position of Gauss.

A More Accurate Verification of the Inverse Square Law.—The expressions obtained for the tangent A and B positions have been derived on the assumption that the inverse square law is true. If, therefore, the values of $\frac{M}{H_0}$ obtained by using a given magnet in the two positions are consistent, the inverse square law will have been verified. Unfortunately, however, the uncertain factor in these equations is the value to be assigned to l, the semi-length of the magnet. This difficulty may be avoided by using a very short magnet, so that $\frac{l}{r} \to 0$, and by measuring the deflexions with the sensitive magnetometer just described. When $\frac{l}{r} \to 0$ we have

$$\frac{M}{H_0} = \tfrac{1}{2}r^3 \tan\theta_A, \text{ and } \frac{M}{H_0} = r^3 \tan\theta_B.$$

Thus
$$\frac{\tan\theta_A}{\tan\theta_B} = 2.$$

If the law of attraction were one of the inverse n-th power we should have

$$\frac{\tan\theta_A}{\tan\theta_B} = n\,.$$

for the intensity at a point on the axis of a bar magnet would be

$$H_A = m\left\{\frac{1}{(r-l)^n} - \frac{1}{(r+l)^n}\right\} = \frac{m}{r^n}\left\{\frac{1}{\left(1-\frac{l}{r}\right)^n} - \frac{1}{\left(1+\frac{l}{r}\right)^n}\right\}$$

$$= \frac{m}{r^n}\left[1 + n\,.\,\frac{l}{r} - 1 + n\,.\,\frac{l}{r}\right], \text{ when } \frac{l^2}{r^2} \text{ and higher terms}$$

are neglected,

$$\frac{nM}{r^{n+1}}.$$

Similarly, for H_B we should have [cf. Fig. 40·4],

$$H_B = \frac{2m}{BN^n}\cos\alpha = \frac{2m}{(r^2+l^2)^{\frac{n}{2}}}\cdot\frac{l}{(r^2+l^2)^{\frac{1}{2}}} = \frac{M}{(r^2+l^2)^{\frac{1}{2}(n+1)}}$$

$$= \frac{M}{r^{n+1}}, \text{ if } \frac{l}{r} \to 0.$$

Consequently $\dfrac{M}{H_0}$ would equal $\dfrac{1}{n}r^{n+1}\tan\theta_A$ and $r^{n+1}\tan\theta_B$ respectively for the two positions; hence $\tan\theta_A = n\tan\theta_B$.

In 1832 Gauss carried out a series of experiments on the above lines. The magnet which he used was about a foot long and had a mass of almost 1 lb. The magnetometer was of the reflecting type and carried a plane mirror at one end of the suspended needle. A horizontal scale in mm., etc., and more than a metre long, was fixed above a telescope 5 metres away from the magnetometer needle. The telescope was arranged so that the image of the horizontal scale above it, which was formed by reflexion in the plane mirror, could be observed. In this way small angular displacements of the magnetometer needle could be measured as explained previously [cf. p. 378]. A displacement of one scale division across the fiducial mark in the observing telescope corresponded to an angular rotation of 22 seconds. A rotation of one-tenth this amount could be estimated by eye. Now Gauss was fully appreciative of the fact that the chief uncertainty in these experiments was in estimating the distance between the poles of the magnet. He therefore decided to work with the magnet at distances between 1 and 4 metres from the magnetometer so that the fraction $\dfrac{l}{r}$ was always small and certainly negligible when r exceeded 2 metres.

Some of the results he obtained on 24–28 June 1832 are included in the following table.

Distance of magnet from magnetometer (r)	Deflexion of needle	
	Tan A position	Tan B position
1·1 metre.	2° 13′ 51·2″	1° 10′ 19·3″
2·0 ,,	37′ 6·2″	19′ 1·6″
4·0 ,,	4′ 35·9″	2′ 22·2″

Since $\tan\theta_A \simeq 2.\tan\theta_B$, it was verified that n was equal to 2 within the limits of experimental error. Now it would be very remarkable if such a universal law should contain an index 2 plus

or minus a very small fraction. It is therefore concluded that the value of n is exactly 2.

It is interesting to note that with the same apparatus Gauss made the first absolute determination of H_0, the horizontal component of the earth's magnetic field. Until then, values for H_0 at different places had only been compared by an oscillation method [cf. p. 771].

The Comparison of Magnetic Moments and the Adjustments of a Deflexion Magnetometer.—Magnetic moments may be compared with the aid of a deflexion magnetometer. The magnetometer is first made level and then arranged so that the pointer attached to its needle sets at the zero marks on the circular scale inside the instrument. The scale in cm., etc., used to measure the distance of the centre of any magnet from the centre of the needle is then placed parallel, or at right angles, to the pointer as desired. The following procedure is adopted irrespective of whether the tangent A or the tangent B position of Gauss is being used. Fig. 40·9 indicates the positions of the magnet when the magnetometer is in the A position of Gauss.

(i) The magnet is placed with its centre at the desired distance away from the needle—Fig. 40·9 (*a*). After gently tapping the case of the magnetometer to overcome the effects of any sticking at the pivot, on which the needle rotates, the positions of both ends of the pointer on the circular scale are noted. In this way any error due to the fact that ***the axis of rotation of the needle may not pass through the centre of the circular scale*** is eliminated if the departure from the ideal conditions is not large.

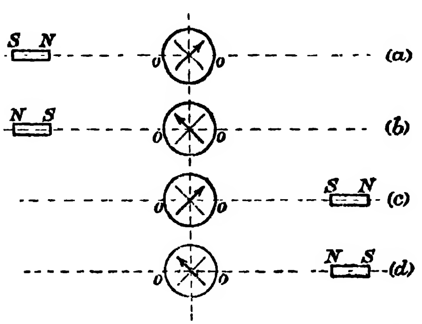


FIG. 40·9.—Comparison of Magnetic Moments and the adjustment of a Deflexion Magnetometer.

(ii) The magnet is then turned end for end and the positions of the ends of the pointer again noted —Fig. 40·9 (*b*). If ***the magnet is not magnetized symmetrically*** the effects due to this are eliminated by proceeding in this way.

(iii) The observations are then repeated with the magnet at the same distance on the other side of the magnetometer needle—Fig. 40·9 (*c*) and (*d*). Any error arising from the fact that ***the needle may not be pivoted at the centre of the graduated arm*** are thereby eliminated.

The mean of the eight readings thus obtained will be equal to that deflexion which would be obtained if the settings of the scales with respect to the needle and each other were ideal, provided that in no instance does the mean differ very much from any one of the eight above readings. If, on any occasion, a large difference should be found, it probably means that the pointer is not at right angles to the axis of the magnetometer needle.

Experiment 1.—Each magnet is placed in an end-on position and the corresponding deflexions determined. If ***suffixes*** refer to the two magnets we have

$$\frac{M_1}{H_0} = \frac{(r_1^2 - l_1^2)^2}{2r_1} \tan\theta_1, \text{ and } \frac{M_2}{H_0} = \frac{(r_2^2 - l_2^2)^2}{2r_2} \tan\theta_2.$$

Hence $$\frac{M_1}{M_2} = \left[\frac{r_1^2 - l_1^2}{r_2^2 - l_2^2}\right]^2 \cdot \frac{r_2}{r_1} \cdot \frac{\tan\theta_1}{\tan\theta_2}.$$

Instead of determining the deflexions due to each magnet the position of the second magnet may be adjusted until the two deflexions are equal when the above equation becomes

$$\frac{M_1}{M_2} = \left[\frac{r_1^2 - l_1^2}{r_2^2 - l_2^2}\right]^2 \cdot \frac{r_2}{r_1}.$$

These experiments may be repeated with the magnets in the broadside-on position.

Experiment 2.—The measurement of angles may be eliminated by using the following ***null method.*** The two magnets are placed on opposite sides of the magnetometer and the position of one of them adjusted until the needle is not deflected from its zero position. Under these conditions the intensities at the centre of the needle due to each magnet separately must be equal so that

$$\frac{2M_1r_1}{(r_1^2 - l_1^2)^2} = \frac{2M_2r_2}{(r_2^2 - l_2^2)^2},$$

or, $$\frac{M_1}{M_2} = \left[\frac{r_1^2 - l_1^2}{r_2^2 - l_2^2}\right]^2 \cdot \frac{r_2}{r_1}.$$

Moment of Inertia of a Rigid Body about an Axis of Rotation.—Suppose that a rigid body is rotating about a fixed axis with angular velocity ω. Consider a portion of that body, so small that it may be regarded as a material particle. Let m be its mass and r its least distance from the axis of rotation. Then the linear velocity of that particle is $v = r\omega$. Its kinetic energy is

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}mr^2\omega^2.$$

For the whole body, the kinetic energy will be

$$\Sigma\tfrac{1}{2}mr^2\omega^2 = \tfrac{1}{2}\omega^2\Sigma mr^2,$$

where the summation refers to all the particles which constitute the

rigid body. The quantity Σmr^2 is termed the ***moment of inertia of the body about the particular axis of rotation considered.***

The Vibration Magnetometer.—When a magnet oscillates freely in a horizontal plane in a uniform magnetic field the motion is simple harmonic if the motion is restricted so that the amplitude is small and there is no couple due to torsion in the supporting filament, or to friction at the pivot. The periodic time in seconds is expressed by

$$T = 2\pi\sqrt{\frac{I}{MH}},$$

where M is the magnetic moment of the magnet, H the horizontal component of the magnetic field [generally the earth's], and I the moment of inertia of the magnet about its axis of rotation. For a given magnet this is a constant depending on its mass, shape, and the axis about which it oscillates. For a rectangular bar of mass m, of length a and breadth b, oscillating about an axis through its centre of gravity and normal to the plane containing a and b,

$$I = m\left[\frac{a^2 + b^2}{12}\right].$$

For a cylindrical magnet of mass m of total length $2a$ and radius r performing oscillations about an axis through its centre of gravity and normal to its length

$$I = m\left[\frac{a^2}{3} + \frac{r^2}{4}\right].$$

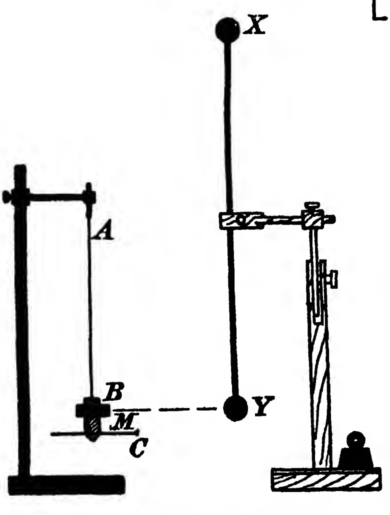


Fig. 40·10.—Searle's Vibration Magnetometer.

Searle's Magnetometer.—This consists essentially of a fine thread of unspun silk, AB, supported at its upper end, and carrying at its lower end a brass cylinder tapering to a point as shown in Fig. 40·10. This point enables the position of the central axis of the block to be determined. The brass block carries a short magnet, M, arranged horizontally. A light aluminium pointer, C, about 10 cm. long enables the oscillations to be observed more easily. The brass block serves to increase the moment of inertia of the system about its axis of rotation so that

its period becomes slow enough for accurate observations to be obtained. Unspun silk is used for suspending the magnet, since the effect of torsion in this material is negligible.

With the help of this apparatus a magnetic survey of the laboratory may be made. The equation for the period of an oscillating magnet may be written $HT^2 = \kappa$, where κ is a constant. If H is the horizontal field at some point, and this is known, the value of κ may be calculated when T is known. The value of H at other points may be deduced from the value of κ thus obtained and the observed time of swing at the point in question.

Comparison of Two Horizontal Magnetic Fields.—Two horizontal magnetic fields could, in general, be compared by the above method if it were possible to isolate them, but as a rule the needle will oscillate in a field which is the resultant of one of the fields to be compared and the earth's horizontal field, H_0. To compare the two given fields it is therefore customary to arrange them so that their directions coincide with that of H_0 and then make the following observations:—If H_1 and H_2 are the fields, H_0 is the earth's horizontal field, and the times of oscillation of the needle are T_1 and T_2 when the two fields are arranged parallel to H_0 and such that the composite fields are $(H_1 + H_0)$ and $(H_2 + H_0)$, we have

$$(H_1 + H_0)T_1^2 = 4\pi^2 IM^{-1} = \kappa \text{ (say)}$$

But $\kappa = H_0T_0^2$, where T_0 is the period of oscillation in the earth's magnetic field H_0.

Hence $H_1T_1^2 = H_0(T_0^2 - T_1^2)$.

Similarly $H_2T_2^2 = H_0(T_0^2 - T_2^2)$

$$\therefore \frac{H_1}{H_2} = \frac{\left(\frac{T_0}{T_1}\right)^2 - 1}{\left(\frac{T_0}{T_2}\right)^2 - 1}.$$

Oscillation Method for Verifying the Inverse Square Law.—A ball-ended magnet, XY, Fig. 40·10, is supported with its axis vertical and its lower pole in the horizontal plane containing the needle of a Searle magnetometer. If the magnet is long compared with the distance from the centre of the lower sphere to the centre of the oscillating needle, the effect of the upper pole may be neglected. The polarity of the lower sphere should preferably be such that the horizontal magnetic field at the centre of the needle is increased. This condition is easily tested, for if it exists the period of the needle will be shortened. Let r be the distance of the lower pole from the centre of the needle when the period is T and the total horizontal field $(H + H_0)$, where H is the contribution due to the lower magnetic pole, and H_0 is due to the earth.

Now $H + H_0 = \frac{\kappa}{T^2}$ and $H_0 = \frac{\kappa}{T_0^2}$.

If we assume an inverse n-th power law, $H = \frac{m}{r^n}$, where m is the pole strength of the magnet. Hence

$$\frac{H}{H_0} + 1 = \left(\frac{T_0}{T}\right)^2,$$

or

$$\left(\frac{T_0}{T}\right)^2 - 1 = \frac{H}{H_0} = \frac{m}{H_0 r^n}.$$

$$\therefore n \log r = \log C - \log \left[\left(\frac{T_0}{T}\right)^2 - 1\right],$$

where $C = \frac{m}{H_0}$, a constant.

A series of observations should therefore be made and the graph

$$x = \log r,\ y = \log\left[\left(\frac{T_0}{T}\right)^2 - 1\right],$$

constructed. This should be a straight line whose slope is $-n$. It will be found that $n = 2$.

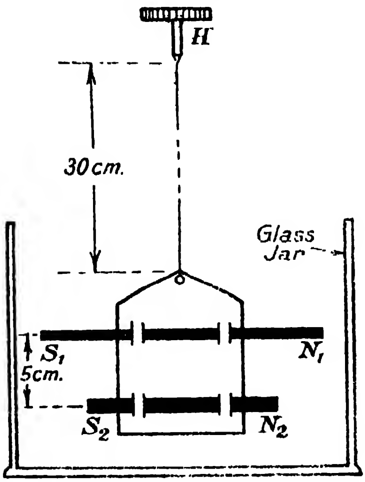


FIG. 40·11.—Comparison of Magnetic Moments by an Oscillation Method.

Comparison of Magnetic Moments by Oscillation Methods.—The two magnets are suspended, in turn, by means of unspun silk, so that they perform oscillations in the earth's horizontal magnetic field H_0 and about a vertical axis passing through their centres of gravity. Their periodic times T_1 and T_2, where the suffixes refer to the first and second magnets, having been determined, we have

$$T_1 = 2\pi\sqrt{\frac{I_1}{M_1 H_0}}, \text{ and } T_2 = 2\pi\sqrt{\frac{I_2}{M_2 H_0}}.$$

Consequently

$$\frac{M_1}{M_2} = \frac{I_1 T_2^2}{I_2 T_1^2}.$$

The objection to this method is that its calculation involves a knowledge of I_1 and I_2. In the following method such knowledge is not necessary.

The two magnets are supported horizontally by passing them

through slots cut in a sheet of thick paper. The whole is suspended by a piece of unspun silk—cf. Fig. 40·11. First, let the north poles of each magnet point to the magnetic north. Let T_1 be the period of vibration for small oscillations. Then reverse one of the magnets—let T_2 be the period. The moment of inertia of the system about a vertical axis through its centre of gravity is not altered if care is taken that the axis of rotation always passes through the centres of both magnets, but the total magnetic moment is $M_1 + M_2$, and $M_1 - M_2$ in the two instances respectively. If the restoring couple due to torsion in the fibre is negligible, we have,

$$T_1 = 2\pi\sqrt{\frac{(I_1 + I_2)}{(M_1 + M_2)H_0}}, \text{ and } T_2 = 2\pi\sqrt{\frac{I_1 + I_2}{(M_1 - M_2)H_0}}.$$

Hence

$$\frac{M_1 - M_2}{M_1 + M_2} = \frac{T_1^2}{T_2^2},$$

or,

$$\frac{M_1}{M_2} = \frac{T_1^2 + T_2^2}{T_2^2 - T_1^2}.$$

In this experiment it is important that the magnets should be as far apart as possible in order to diminish the strength of the induced poles, and hence their effect. The effects of air currents on the motion are eliminated by surrounding the magnets by a glass jar.

In practice, only small magnets may be used in these oscillation experiments since it is difficult to find a suspension sufficiently strong to support the weight of the system, and yet not exert a restoring couple on it.

So far it has been assumed that the torsion in the suspension is negligible. It may happen, however, that one end of the suspension has been twisted through a large angle relatively to the other—the torsion couple may be large under such circumstances. To free the system from such a couple, the magnets are replaced by brass rods, and the system allowed to come to rest. The head, H, carrying the silk is then rotated until the paper lies in the magnetic meridian. The torsional couple is then very small. When the magnets are reinserted the system, even when it oscillates, will be free from a large torsional couple.

EXAMPLES XL

1.—Calculate the field strength at a point on the axis of a bar magnet of pole strength 100 units and magnetic length 10 cm. The point is 45 cm. from the centre of the magnet.

2.—ABC is a triangle right angled at B. At A and B north-seeking poles of strengths 16 and 30 units respectively are placed. If AB = 20 cm. and BC = 15 cm., calculate the magnetic field strength at B.

3.—A bar magnet measures 20 cm. × 2 cm. × 3 cm. The intensity of magnetization in the magnet is 6·2 gauss. Calculate the pole strength, and magnetic moment, of the magnet.

4.—A magnet of moment 81·4 erg.oersted.$^{-1}$ is suspended in the meridian and then deflected through 41°. What is the couple acting upon it if H_0 = 0·182 oersted?

5.—A magnet makes 10 complete swings in 84 sec. at a point where H_0 = 0·20 oersted. Find the time of swing when H_0 = 0·26 oersted.

6.—Two magnets of the same material and size make 50 swings in 6 min. 18 sec. and 6 min. 43 sec. at the same station. If the first magnet has a moment 84 units, calculate that of the second.

7.—A compass needle having a magnetic moment 850 erg.oersted.$^{-1}$ is rotated through an angle of 55°. Calculate the couple necessary to maintain the needle in this position and the work done in rotating the needle from its position of rest. [H_0 = 0·18 oersted.]

8.—How would you compare the strengths of two uniform magnetic fields superposed at right angles to each other? How would you compare them if the two fields were entirely separate?

9.—Derive an expression for the intensity of the magnetic field at any point on the prolongation of the axis of a bar magnet. Explain how the expression may be used in the experimental comparison of the magnetic moments of magnets.

10.—Deduce expressions for the magnetic field strength due to a bar magnet in the tangent A (end-on) and tangent B (broadside-on) positions of Gauss respectively. Explain the units in which magnetic field strength is measured.

11.—Describe how you would compare the magnetic moments of two magnets of the same size and shape, (*a*) using a deflexion magnetometer, (*b*) by a vibration method.

12.—Describe and explain how you would compare the magnetic moments of two short magnets by using a deflexion magnetometer.

13.—A compass needle is set swinging in a magnetic field. What factors determine its period of oscillation? Describe experiments you would make to illustrate your answer.

14.—Explain how it is that bar magnets of different sizes and shapes may have equal magnetic moments. How could you find which of two given bar magnets has the greater pole strength?

15.—Explain *magnetic moment, moment of inertia.*

A bar magnet is placed on a horizontal table and a neutral point in its field is located. A small magnet suspended by a long silk thread is placed with its centre immediately above the neutral point. The bar magnet is then reversed, end for end, and the small magnet is found to make twelve complete oscillations per minute. How many oscillations will it make per minute when the bar magnet is removed?

16.—Explain what is meant by the statement '$H_0 = 0{\cdot}20$ *oersted.*' A bar magnet 20 cm. long stands upright with its north pole resting on a table. Give a diagram showing the general distribution of the lines of magnetic force in the plane of the table. If there is a neutral point 6 cm. from the magnet, calculate the magnetic moment of the magnet.

17.—Define the terms *unit pole, magnetic moment.*

A horse-shoe magnet having its poles 5 cm. apart is set so that the line joining them lies E.–W. and passes through the centre of a deflexion magnetometer distant 10 cm. from the nearer pole. If the deflexion of the needle of the magnetometer is 35° estimate the pole strength of the magnet. Assume H to be 0·18 oersted.

CHAPTER XLI

TERRESTRIAL AND SOLAR MAGNETISM

The Magnetic Field Round the Earth.—Round the earth there is a magnetic field, the intensity of which varies from place to place, and to a less extent daily and yearly. DR. GILBERT believed that the earth was a large magnet with its poles at opposite ends of a diameter of the earth. Of course there is no actual magnet there; in fact, the origin of the earth's magnetism is a

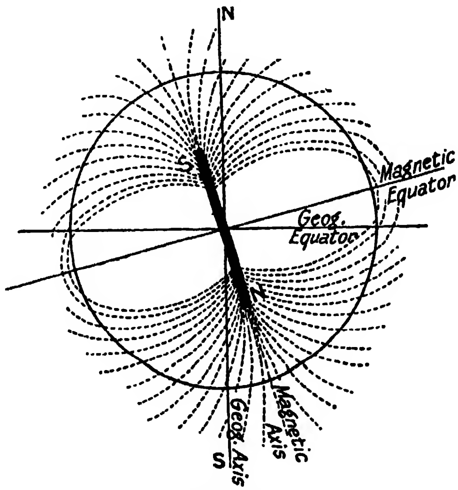
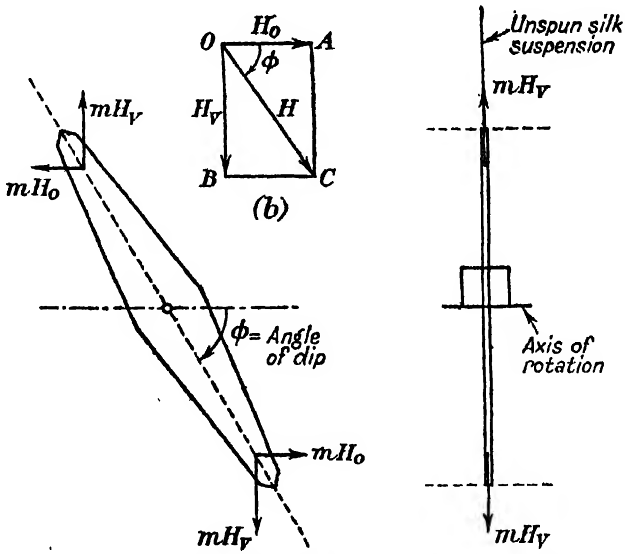


FIG. 41·1.

mystery, but the behaviour of the earth's magnetism is as if a powerful magnet were present at its centre with its axis pointing approximately south and north—such a hypothetical magnet is shown in Fig. 41·1. It will be noticed that the magnetic axis and equator do not coincide with the corresponding geographical positions, and that the hypothetical magnet has south-seeking magnetism at the pole which points towards the geographical north. Similar remarks apply to the southern hemisphere.

The Earth's Magnetic Elements.—If a magnet is suspended freely as in Fig. 41·2 (*a*) it is, in general, inclined to the horizontal.

The magnet sets itself so that its magnetic axis lies along the direction in which the earth's magnetic intensity acts, i.e. in the direction OC, Fig. 41·2 (*b*). The angle ϕ of the diagram is called the ***angle of dip.*** Now the total magnetic intensity IH (or H_R), which is represented in magnitude and direction by OC may be resolved into two components represented by OA and OB respectively. These are termed the horizontal and vertical components of the earth's magnetic field, and are referred to as H_0 and H_V respectively; in fact, H_0 is the magnetic field with which another field is generally compared in magnetometer experiments.



(*a*) Side-View of Needle (*c*) Edge-View of Needle

FIG. 41·2.

If ϕ is the angle of dip, then

$$\frac{H_V}{H_0} = \frac{OB}{OA} = \frac{AC}{OA} = \tan\phi.$$

Similarly,

$$\frac{H_0}{H} = \frac{OA}{OC} = \cos\phi, \text{ or } H_0 = H\cos\phi,$$

and

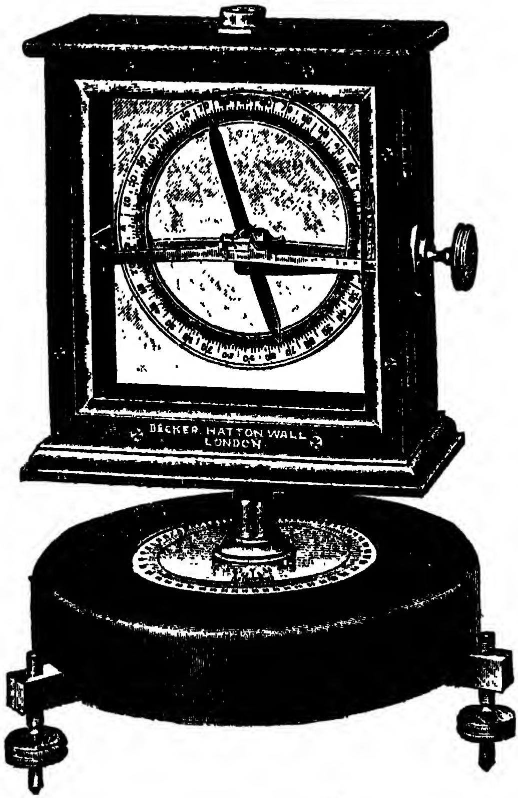
$$\frac{H_V}{H} = \frac{AC}{OC} = \sin\phi, \text{ or } H_V = H\sin\phi.$$

In an earlier chapter it has been stated that the axis of a suspended magnet only points approximately to the geographical north and

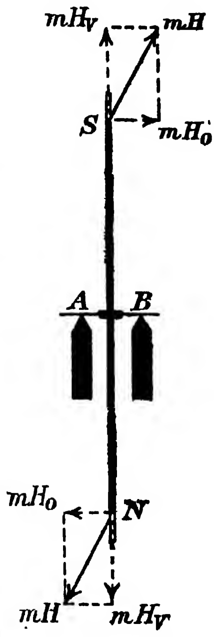
south. The vertical plane passing through the axis of such a magnet is called the ***magnetic meridian,*** as distinct from the geographical meridian, which is the vertical plane passing through a line of longitude. The angle between these two planes is called the ***angle of declination.***

When, at any station, the declination, dip, and the horizontal component of the earth's magnetic field, are known, the magnetic field at that station is completely defined.

Measurement of the Angle of Dip.—A simple model of an instrument used for the determination of the angle of dip is shown in Fig. 41·3 (*a*). Such an instrument is called a ***dip circle.*** It con-



(*a*).—Dip Circle.



(*b*).—Forces on a Dip Needle when its axis of rotation is in the Magnetic Meridian.

FIG. 41·3.

sists essentially of a magnetized needle capable of rotation about a horizontal axis, the points of support being agate knife-edges. These should be kept free from grease. When the instrument is

not in use the needle is raised from its position of rest on the knife-edges by means of two sliding pieces with V-shaped grooves. These sliding pieces move together, their motion being controlled by means of a screw-head outside the case of the instrument. The positions of the ends of the needle are given by a vertical circular scale graduated in degrees, etc. The instrument is protected from dust [which causes the needle to stick on the knife-edges], and from draughts, by means of a case, the front and back of which are glass plates.

To use the instrument, all pieces of iron having been removed from the immediate vicinity, it is first levelled by means of the screws supporting the base. The upper box, capable of rotation about a vertical axis, is turned about that axis until the needle is vertical; in this position the effect of the horizontal component of the earth's field is nullified, for otherwise the needle would not be vertical. The plane of the needle is then normal to the magnetic meridian. For consider the forces acting on the poles of the needle—cf. Fig. 41·3 (*b*). They may be resolved into rectangular components as shown. The horizontal forces mH constitute a couple, but they cannot cause the needle to rotate since it is supported at A and B. The vertical components constitute a couple whose moment is not zero unless the needle is vertical. The needle, therefore, sets with its axis vertical. The case is then rotated through 90°, when the needle is in the magnetic meridian, and the angle of dip is observed on the circular scale; the positions of both ends of the needle are recorded.

Several errors arise in using a dip circle: their effects may be eliminated as follows:

(i) ***The axis of rotation of the needle may not pass through the centre of the vertical scale.*** The effect of this is eliminated by observing each end of the needle and using the mean of the apparent angles of dip. For if O_1 and O_2, Fig. 41·4 (*a*), are the centres of the scale and the point in which the axis of rotation cuts the needle, respectively, then the actual readings are really measures of the angles S_2O_1Y and N_2O_1X respectively. If N_1S_1 is drawn through O_1 parallel to the needle, N_2S_2, the true dip is $N_1\hat{O}_1X$ or $S_1\hat{O}_1Y$—say ϕ. Now $N_2\hat{O}_1X$ is greater than ϕ by $N_1\hat{O}_1N_2$, and $S_2\hat{O}_1Y$ is less than ϕ by $S_1\hat{O}_1S_2$. Since $N_1N_2 = S_1S_2$, the above differences are equal, so that ϕ is the mean of the observed readings.

(ii) ***The centre of gravity of the needle may not lie on the axis of rotation.*** Suppose that AB, Fig. 41·4 (*b*), is the needle and that its centre of gravity G does not lie on the axis of rotation which passes through O: this axis is normal to the plane of the diagram. Now the moment about O of the weight W of the needle is equal to the sum of the moments of two equal weights W at P and Q respectively. Thus the lack of coincidence of G with O

may be considered in two parts, (*a*) that due to the displacement of G in a direction normal to AB, and (*b*) that due to its displacement parallel to AB.

If the end A is dipping, i.e. the north pole of the magnet is near to A if the needle is used in the northern hemisphere, and the needle is placed so that Q is above O, then the effect will be that the measured angle of dip will be too large: if the needle is reversed on its bearings so that Q is below O, then the effect will be that the measured angle of dip is too small by an equal amount. By taking the mean of the four readings so far obtained under (i) and (ii) the errors due to the axis of rotation of the needle not passing through the centre of the vertical scale and to the displacement of G at right angles to AB are eliminated.

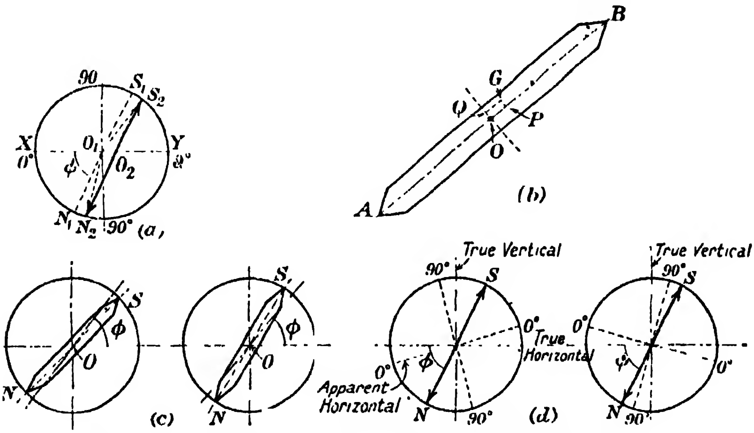


FIG. 41·4.—Errors in a Dip Circle—Their Elimination.

The procedure under (ii) also eliminates another error at the same time. We shall consider this next and later return to the question of the error arising from the displacement of G parallel to AB. The error so eliminated is that which arises from the fact that:

(iii) ***The magnetic axis of the needle may not coincide with its geometrical axis.*** The elimination of the error arising from this cause is effected by reversing the needle relatively to the vertical scale by removing it from its bearings, turning it back to front, and replacing it—cf. Fig. 41·4 (*c*). This has already been done under (ii).

(iv) ***The zero line of the vertical scale may not be perfectly horizontal.*** The effect of this is eliminated by rotating the instrument through 180° at each stage of the process. The needle is still in the magnetic meridian, but the 'apparent horizontal'—the

line joining the zeros on the vertical scale—will now be tilting in the opposite direction: cf. Fig. 41·4 (*d*). The positions of the ends of the needle are again observed. The mean gives a value for the dip corrected for this error.

(v) When the eight readings required under (i), (ii), (iii) and (iv) have been made the only source of error outstanding is that due to the displacement of G parallel to AB. As long as A is the end of the needle which dips, the effect of this displacement parallel to AB will be to make the measured angle of dip too small, but if the needle is remagnetized so that the end B dips, i.e. the polarity of the needle is reversed, then the above displacement of G will make the measured angle of dip too large. Thus eight more observations are required and the mean value of the complete sixteen readings is one from which errors attributable to the above causes have been eliminated if the discrepancies are small.

The Angle of Declination.—Since this is the angle between the geographical and magnetic meridians it is first necessary to locate the geographical meridian. This may be done by observing the direction in which the shadow of a vertical string lies when the sun is in the geographical south—it must not be assumed that the sun is in the south at noon on all occasions. The exact time when this position is reached can always be ascertained from a nautical almanac.

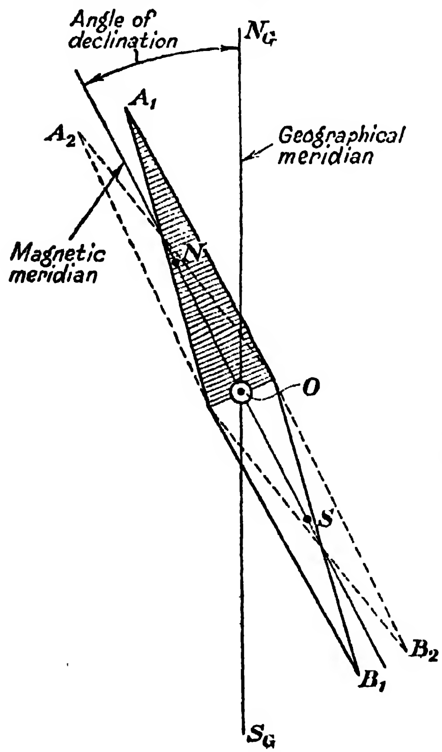


FIG. 41·5.—Determination of the Magnetic Meridian.

It then remains to determine the magnetic meridian. If it were possible to obtain a magnetized needle with its magnetic axis coinciding exactly with its geometrical axis then the measurement could be made easily. Unfortunately, this ideal cannot be realized, and so the following method is adopted.

The magnetic needle is suspended by a piece of unspun silk so that it may swing in a horizontal plane. In general, the needle will come to rest with its geometric axis AB inclined to the geographical meridian N_GS_G, i.e. it will be in the position A_1B_1, Fig. 41·5. Then the angle observed is that between the geographical meridian and the geometrical axis of the needle. To determine the necessary correction the needle is inverted face for face and then allowed to come to rest. In each instance the needle's magnetic axis NS will lie in the magnetic meridian (this assumes a torsionless suspension) and the geometrical axis will take up symmetrical positions with respect to the magnetic meridian. If A_1B_1 and A_2B_2 are these positions, the bisector of the angle A_1OA_2 (or B_1OB_2) gives the magnetic meridian.

[The experiment may be made using a bar magnet or even a magnetized circular disc of iron. The chosen magnet then has two straight pieces of copper wire fastened to its extremities with a little soft wax to define a fiducial line in the magnet, and the whole is supported in a stirrup by means of a silk thread. Immediately below the magnet is placed a sheet of white paper. When the magnet has come to rest, pencil marks A_1 and B_1 are made to indicate the positions of the 'pins.' The magnet is then placed with its lower side uppermost and the experiment repeated—the points A_2 and B_2 are thus found.]

Experimental Determination of H_0.—To determine the absolute value of the horizontal component of the earth's magnetic field two experiments are necessary. In the first or deflexion experiment the value of $\frac{M}{H_0}$ is ascertained by using a magnet in the tangent A [or B] position of Gauss. If θ is the mean deflexion of the magnetometer needle, where r and l have their usual significance, we have,

$$\frac{M}{H_0} = \frac{(r^2 - l^2)^2}{2r} \tan \theta = \alpha \text{ erg.oersted.}^{-2} \text{ [say]} \quad . \quad . \quad \text{(i)}$$

The second or oscillation experiment consists of a determination of MH_0 by suspending the given magnet in the earth's field and observing its period T. Then, with the notation already explained,

$$T = 2\pi \sqrt{\frac{I}{MH_0}}, \text{ or } MH_0 = \frac{4\pi^2 I}{T^2} = \beta \text{ erg. (say)} \quad . \quad \text{(ii)}$$

From the quotient of equations (i) and (ii) we have,

$$H_0^2 = \frac{\beta}{\alpha}, \text{ or } H_0 = \frac{2\pi}{T(r^2 - l^2)} \sqrt{\frac{2rI}{\tan \theta}} \text{ oersted.} \quad . \quad . \quad \text{(iii)}$$

[If the tangent B position is used, the final equation reduces to

$$H_0 = \frac{2\pi}{T} \sqrt{\frac{I}{(r^2 + l^2)^{\frac{3}{2}} \tan \theta}} \text{ oersted.]}$$

These same two experiments also enable us to determine the absolute value of M, for the square root of the product of the first two equations gives

$$M = \sqrt{\alpha\beta} = \frac{\pi(r^2 - l^2)}{T}\sqrt{\frac{2 . I}{r}\tan\theta}\ \text{erg.oersted.}^{-1} \quad . \quad \text{(iv)}$$

The chief uncertainty in this experiment lies in an assumed value for $2l$, the magnetic length of the magnet. Since, however, equation (i) may be written

$$\left(\frac{2M}{H_0}\right)^{\frac{1}{2}} (r \cot\theta)^{\frac{1}{2}} = r^2 - l^2,$$

it follows that if a series of corresponding values of r and θ is obtained and the straight line $x = (r \cot \theta)^{\frac{1}{2}}$, $y = r^2$, obtained by plotting, the value of $\frac{M}{H_0}$ may be deduced from the slope of the line without reference to a value for l: in fact l^2 is the intercept on the y-axis, but a knowledge of its value is not essential.

The Kew-pattern Unifilar Magnetometer.—By means of this instrument reliable values of the declination and of the horizontal component of the earth's magnetic field may be determined. The arrangement of this instrument for determining the declination and for observing the time of oscillation of the magnet in the vibration part of the experiment for finding H_0 is indicated in Fig. 41·6. The magnet consists of a hollow steel cylinder, A, fixed to a brass collar. This collar also carries a hollow brass cylinder the purpose of which is explained later. A scale S, graduated in mm., is fixed at one end of A while L is a convex lens arranged at the other end. The distance between S and L is equal to the focal length of the lens so that a beam of parallel light emerges from L. This light is received by a telescope F focused for parallel light, and D is a plane mirror used to illuminate the scale by reflected light The magnet and its accessories are suspended from a torsion head K by means of an unspun silk fibre. The whole is mounted inside a box provided with suitable windows. The telescope F is attached to an arm (as in a spectrometer) and is capable of rotation about a vertical axis. N is a circular scale used to determine the position of F.

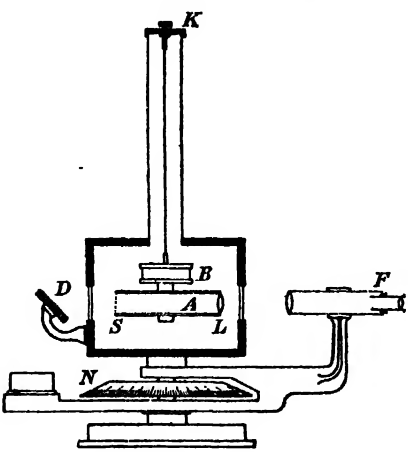


FIG. 41·6.—Kew-pattern Unifilar Magnetometer.

One of the greatest troubles in an accurate determination of the

declination is to free the suspension from torsion. The residual torsion is reduced almost to zero by replacing the magnet A by a brass plummet of about the same mass and allowing this to swing until it comes to rest. Since the material of the plummet is non-magnetic, the position of rest will be such that the suspension is free from torsion. The torsion head is then rotated so that the rest-position of the plummet is in the magnetic meridian. When the plummet is removed and the magnet A replaced the suspension will be practically free from torsion when the magnet is in the magnetic meridian. The effects of the rigidity of the material of the suspension are minimized by using unspun silk (these effects only come into play when the magnet swings).

The eye-piece of the observing telescope is provided with vertical and horizontal cross-wires, and the magnet A is adjusted so that the divisions on the scale S are vertical. The telescope is rotated until the image of the central division on S (the zero) appears to coincide with the vertical cross-wire in the telescope. The final adjustment of the position of the telescope is made by means of a slow-motion screw. Since it is difficult to bring the magnet absolutely to rest it is more usual to adjust the position of the telescope until the apparent angle of swing of the magnet is bisected by the vertical wire in F. The position of the telescope on N is noted and the observations are repeated with the magnet A rotated 180° about a horizontal axis so that S is inverted. The mean reading of the positions of the telescope eliminates any error arising from the fact that the axis of magnetization may not coincide with the axis of the optical system.

It now remains to determine the geographical meridian. From Nautical Tables, the latitude and longitude for the station where the observations are being carried out being known, the azimuth of the sun at any instant is determined. By means of D an image of the sun (duly reduced in intensity with the aid of a piece of smoked glass) is reflected into the optical system and the telescope adjusted so that this image crosses the vertical wire in F at some particular instant. The position of the telescope on the scale N is noted. From the above observed time the direction of the sun at the time of the experiment becomes known; the position of the telescope on the scale N when its axis points north and south is deduced. The declination is equal to the difference between this position of the telescope and its mean position in the former part of the experiment.

For success in locating the position of the sun it is essential that the axis of the telescope should be horizontal, that the plane mirror D should rotate about a horizontal axis, and the normal to the surface of D at any point lie in a plane parallel to a vertical plane containing the optical axis of the telescope.

If a series of observations of the declination at a station are to be made at different times, then it is advisable to use a fixed object whose direction with reference to the geographical meridian is known, instead of determining the direction of the latter on each occasion.

To determine H_0 with the above instrument it is necessary to determine the time of swing of the magnet and its moment of inertia about the axis of rotation. The time of swing is found with the aid of an accurate chronometer. The moment of inertia required is not that of the magnet only but that of the magnet and its carriage. This cannot be calculated. It is determined experimentally as follows.

Let T_1 be the period when the magnet and its carriage oscillate in the earth's horizontal field as above. Let I_1 be the moment of inertia

of the system about the axis of rotation. Then place a brass bar of known moment of inertia about the above axis in the tube B provided for this purpose. This cylinder completely fills B, and B has been adjusted so that when the brass cylinder is introduced the magnet A still swings in the same plane. Let T_2 be the period when the total moment of inertia about the axis of suspension has become $I_1 + I_2$. Then

$$T_1 = 2\pi\sqrt{\frac{I_1}{MH_0}}, \text{ and } T_2 = 2\pi\sqrt{\frac{I_1 + I_2}{MH_0}}.$$

$$\therefore \frac{T_2}{T_1} = \sqrt{1 + \frac{I_2}{I_1}}$$

$$\therefore I_1 = I_2\left(\frac{T_1^2}{T_2^2 - T_1^2}\right).$$

so that I_1 becomes known. For a cylinder of mass m, length $2a$, and radius r, $I_2 = m\left(\frac{a^2}{3} + \frac{r^2}{4}\right)$.

It is only necessary to determine I_1 once, since it is a constant for the system and is independent of the magnetic field in which the instrument is situated.

The second part of the experiment consists in determining the angle through which a small magnet is deflected by the magnet A. This magnet is removed and its place taken by a small magnet carrying a plane mirror on its under side. The plane of this mirror is normal to the axis of the magnet. By means of a lamp and scale arranged as on p. 378, the deflexion of the suspended magnet caused by any external field is measured. In the present instance this field is produced by the magnet A situated in the tangent A or tangent B position of Gauss. Let us assume that it is in the former position. A scale in mm. attached to the magnetometer enables the distance between the centres of the two magnets to be determined. A mean value of the deflexion is deduced from a series of observations made as described on p. 768. The ratio M/H_0 is then calculated from the equation

$$\frac{M}{H_0} = \frac{(r^2 - l^2)^2}{2r} \tan\theta.$$

Other methods for the determination of H_0 (and H_V) will be discussed later [cf. Chap. XLIII].

Magnetic Maps.—The earth's magnetic elements vary from place to place and this variation is best shown by means of lines drawn upon a geographical chart. The lines on such a map indicate places at which the magnetic element, which is being considered, has the same value. Lines of equal dip are called ***isoclinic lines,*** whilst those showing the places of equal declination are called ***isogonal lines,*** Fig. 41·7. The particular isogonal lines for which the declination is zero, i.e. where a magnet points to the geographical north, are termed ***agonic lines.*** The line of zero dip is called the ***magnetic equator*** or ***aclinic line,*** while the two points at which the dip is 90° are termed the ***surface magnetic poles.*** Lines

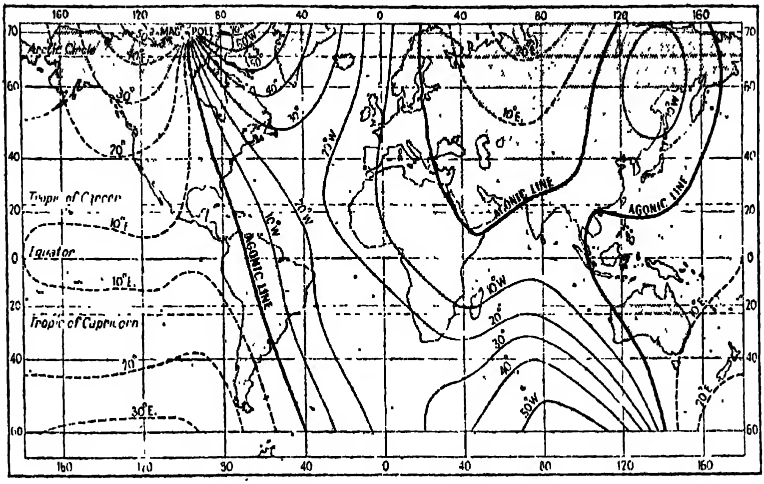


FIG. 41·7 (*a*).—Isogonals for 1942.

Reproduced from British Admiralty Chart No. 3598, with the permission of H.M. Stationery Office and of the Hydrographer of the Navy.

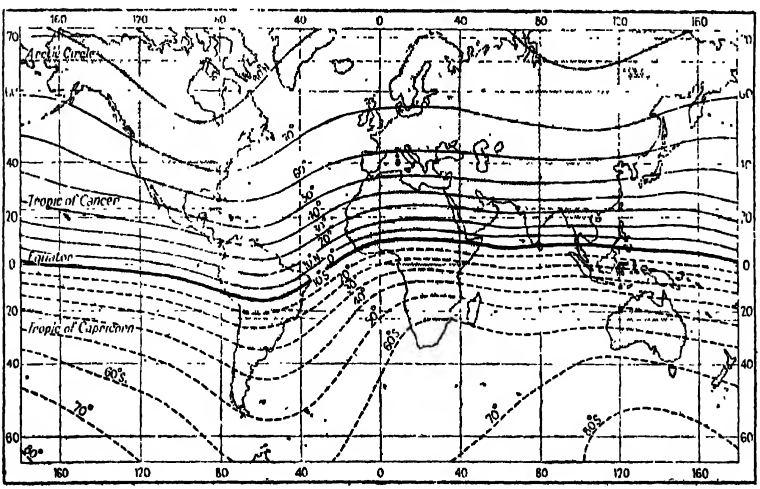


FIG. 41·7 (*b*).—Isoclinics for 1942.

Reproduced from British Admiralty Chart No. 2598, with the permission of H.M. Stationery Office and of the Hydrographer of the Navy.

passing through points having the same value for H are termed ***isodynamic lines.*** The north magnetic pole is situated in North America and was first located by Sir James Ross in 1831 (lat. 73° 31′ N., long. 96° 43′ W.). In 1903 it was situated in latitude 70°40′ N., longitude 60°5′ W. (*Amundsen*). The south magnetic pole was located in 1909 at latitude 72°25′ S., longitude 155°16′ E. (*Scott*). Thus these poles are each about 17° from the geographical poles, but their positions are variable.

A map of the isogonals for the year 1942 is shown in Fig. 41·7 (*a*). It shows that the isogonals converge towards the magnetic poles and that the agonic line passes through America running almost directly from north to south, but that its continuation in the eastern hemisphere is more complicated. A particular feature of this portion of the line is the loop known as the ***Siberian Oval.*** Over that portion of the surface of the globe lying between the two portions of the agonic line and including the Atlantic Ocean, the declination is westerly—also in the Siberian Oval. At other places it is easterly.

The isoclinic lines, or lines of equal dip, are shown in Fig. 41·7 (*b*). These are more regular in their formation than are the isogonals. Each follows a course approximately running from east to west. It will also be noticed that the magnetic equator has a course near to the Equator but actually crosses it once in the Atlantic Ocean and once in the Pacific Ocean.

Continuously Recording Instruments.—Every magnetic observatory, in addition to being equipped with precision instruments for determining the magnetic elements at that station, is also provided

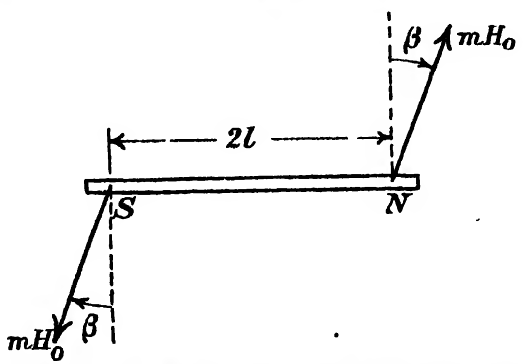


FIG. 41·8.—The Magnet of a Horizontal Variometer.

with three types of instrument recording continuously the local changes in the earth's magnetic field. The first is the ***declination magnetograph.*** The small magnet of this instrument is suspended by a quartz fibre and is attached to the back of a concave mirror. Light falling upon

this mirror is reflected on to a sheet of photographic paper wound on a drum rotating at constant speed. The whole is enclosed in a light tight box and the magnet is surrounded by a large massive copper ring so that the motion of the magnet shall be highly damped [cf. Chap. XLIX]. If the declination were constant a straight line would be found on the paper when developed. Any variation is shown by the excursions of the trace from this line.

Variations in H_0 are detected by the ***horizontal variometer***. This consists of a small magnet arranged as above, but a torsion head to which the suspension is attached is used to twist the magnet into a position at right angles to the direction of the mean magnetic meridian. A bifilar suspension is convenient although the sensitivity of the variometer is somewhat reduced. A plan of the magnet is shown in Fig. 41·8.

Let us first suppose that the direction of H_0 changes by a small amount β, but that H_0 remains constant. The forces in a horizontal

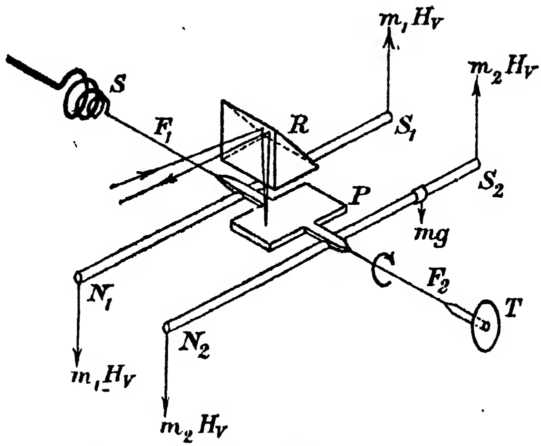


FIG. 41·9.—Watson's Magnetograph.

plane acting on the poles of the magnet are each mH_0, where m is the pole strength of the magnet. These constitute a couple of moment $mH_0 . 2l \cos \beta$. Since β is small this does not differ appreciably from the couple due to the suspension (it is equal to MH_0) and therefore small variations in the direction of H_0 do not affect the instrument.

Now let us suppose that H_0 becomes $(H_0 + \delta H_0)$. Then the moment about a vertical line through the centre of mass of the magnet of the forces acting on it becomes

$$m(H_0 + \delta H_0) . 2l,$$

i.e. the increase is $M . \delta H_0$. The magnet is therefore deflected until the couple due to the suspension is increased to balance the increase in the above moment. Thus the spot of light reflected on to the recording drum moves. This displacement is determined by calibrating the instrument by observing the deflexion caused by placing a magnet whose magnetic moment has been previously determined in a known position with reference to the suspended magnet.

The third type of instrument referred to above is the ***vertical intensity magnetograph.*** None of the instruments designed to record variations in H_V was satisfactory until WATSON constructed the magnetograph described below. In the earlier instruments a magnet was mounted to rotate in the magnetic meridian about a horizontal axis. The magnet was loaded so that it was horizontal when the vertical magnetic field was equal to the mean value of the vertical field at the station in question. The axis of the needle was thus normal to H_V so that any variations in H_V deflected it and could be recorded photographically. Since the needle was supported on a knife edge, mechanical disturbances were a source of much trouble and were only eliminated when Watson constructed the whole of the moving part of the instrument (the magnets excepted) from fused quartz. An additional advantage of his apparatus is that it can be rendered independent of temperature changes: this is most desirable since the magnetic moment of a magnet decreases with rise in temperature.

Watson's magnetograph consists essentially of two magnets N_1S_1 and N_2S_2, Fig. 41·9, of pole strengths m_1 and m_2 respectively. They were rigidly attached to two quartz rods fused to a small quartz plate P, the upper surface of which was polished and flat. The above rods were fused to quartz fibres F_1 and F_2 respectively and these were fixed to a quartz spring S and a torsion head T. The small adjustable mass m is placed in such a position that the ends S_1 and S_2 of the needles which usually point upwards (in northern latitudes) are depressed below the horizontal plane through F_1 and F_2, and the torsion head rotated until the magnets lie in a horizontal plane. Any variation in H_V causes the magnets and the plate attached to them to rotate until the change in the couple acting on the system is balanced by a change in the torsional couple acting on it. The totally internally reflecting prism, R, enables the variations to be detected by a horizontal beam of light incident upon the system in the manner indicated.

Magnetic Storms.—Abrupt changes in the magnetic elements are sometimes reported simultaneously by the different magnetic observatories. These are often associated with the sudden appearance of a large sun spot and a display of the aurora borealis. Changes in the earth's magnetic field are probably due to external influences as the above phenomena suggest.

Variations in Terrestrial Magnetism.—The magnetic field of the earth is constantly changing. The variations are generally slow, and centuries may elapse before the particular magnetic element at any chosen place regains its former value, i.e. the period of the change is very long. These slow-changing variations are termed ***secular changes.*** At the same time the positions of the magnetic poles also change. In addition to these irregular secular changes, very accurate measurements have shown that the magnetic elements also undergo other rapid, but very small, variations. Thus there is a daily period, a lunar month period, a yearly period, a period of 11 years [the spots on the sun have a similar period] and a period of about 26 days. This last time is the period in which the inner core of the sun performs a complete revolution,

for it is a well-known astronomical fact that the sun does not rotate as a rigid body, but that it rotates at different speeds in different latitudes.

Zeeman Effect. Solar Magnetics.—We have already seen how the spectroscope has given us information regarding the elements present in the chromosphere of the sun. The same instrument has also taught us something about the solar magnetic field. In 1895, ZEEMAN, a Dutch physicist, discovered that when the light from a sodium or lithium flame situated in a very intense magnetic field was examined spectroscopically in a direction parallel to the field each line in the usual spectrum became a doublet, whereas if the light was similarly examined in a direction perpendicular to that of the magnetic field then, in addition to every usual line, there were two components associated with it. When direct sunlight is examined by a sensitive spectroscope it is found that doublets occur when the instrument is directed to the centre of the sun, whereas triplets appear if the light examined comes from near the periphery of the sun. This shows that there is a magnetic field of great intensity round the sun and that the field is a radial one.

EXAMPLES XLI

1.—What do you understand by the terms declination, dip, magnetic intensity, H? Describe the use of a dip circle.

2.—A cylindrical magnet of mass 23 gm. makes 10 complete swings in 109 sec when oscillating in the earth's horizontal magnetic field H_0. It is 7·8 cm. long and has a mean diameter of 0·95 cm. When placed with its centre 15 cm. from a magnetometer, the mean deflexion is 42·5°. Calculate a value for M and for H_0.

3.—Write a brief account of the more important properties of the earth's magnetic field.

4.—Define the terms: *magnetic dip, magnetic declination*. Give an account of the method you would adopt to compare the horizontal components of the earth's magnetic field at two points in a laboratory.

5.—The axis about which a dip-needle is movable is slowly rotated in a horizontal plane. Describe and explain the behaviour of the needle during one complete turn of the axis (a) in England, (b) at the magnetic equator. (B.S.S.C. '29.)

CHAPTER XLII

THE MATHEMATICAL THEORY OF MAGNETIC PHENOMENA

Magnetic Media.—Hitherto we have supposed that the magnets whose effects have been studied have been situated in air—more strictly in a vacuum. It is now necessary to consider the changes which occur when the magnet is surrounded by a medium capable of being magnetized itself. Many of the equations derived in this chapter will be obtained from analogies with the corresponding phenomena in dielectrics. Hence, for the present, no attempt will be made to account for the magnetic properties of material media. We shall therefore assume that when an isotropic medium is placed in a magnetic field it acquires a certain magnetic moment per unit volume. This is termed the *intensity of magnetization* in the medium, and is denoted by the symbol J. [This term is discussed more fully in Chap. L.] For most media the direction of J coincides with that of the field.

Magnetic Intensity and Magnetic Induction.—The magnetic intensity at a point in air, [strictly speaking, in a vacuum], has been defined as the force per unit positive pole on a small positive pole placed at the point. When it is desired to measure the force on such a pole inside a piece of iron, or other magnetizable substance, a cavity must first be made in the specimen so that the small pole may be introduced into it. Now the walls of the cavity will exhibit magnetic polarity which will contribute to the total force on the small pole in the cavity. The contribution will be determined, in part at least, by the shape of the cavity, which must therefore be carefully specified if the physical interpretation of this force is to have a definite meaning.

Let us consider the force per unit positive pole on a small positive pole, δm, at the point P, Fig. 42·1 (a), at the centre of a cylindrical cavity, whose diameter is small compared with its length, and whose axis is in the direction of the magnetization at P. The induced magnetism will appear on the ends of this cavity. If J is the intensity of magnetization, and α the cross-section of the cavity, the charges of magnetism at the ends of the cavity will be $J\alpha$ and $-J\alpha$, respectively. If $2l$ is the length of the cylinder, the force on the small pole δm at P due to the magnetism on the walls of the cavity is $\left(\frac{J\alpha}{l^2} + \frac{J\alpha}{l^2}\right)\delta m$. This is zero, since the cavity is very long compared with its width. The force per unit positive pole at P is therefore due to the magnetizing field. Call it H.

Now consider the force on δm when this is at P the centre of a cavity whose length is small compared with its diameter—the cavity resembles a disc—Fig. 42·1 (b). Again let the axis of the cylinder be

parallel to the field. Let β be the area of each plane face of the disc. It is only on these faces that induced magnetism will appear. Now the contribution to the force per unit positive pole on δm due to these induced charges of magnetism is $4\pi J$, a result obtained from analogy with the corresponding problem in electrostatics [cf. p. 709].

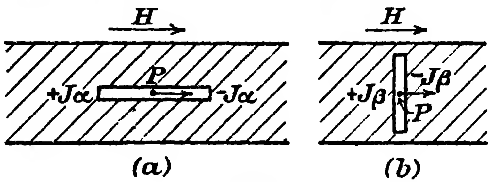


FIG. 42·1.

The actual force per unit positive pole on the small pole in the cavity is obtained by adding together the two quantities H and $4\pi J$. This total force per unit positive pole on the small pole is a measure of the ***magnetic induction***, B, of the material. Hence

$$B = H + 4\pi J.$$

The unit of magnetic induction is the ***gauss***.

Magnetic Susceptibility and Magnetic Permeability.—The quantity χ, defined by the equation, $J = \chi H$, is termed the ***susceptibility*** of the material of the specimen.

The *permeability*, μ, of the medium is defined by the equation $B = \mu H$. Since $B = H + 4\pi J$, it follows that

$$\mu = 1 + 4\pi\chi.$$

Since the magnetic induction B is related to the magnetic intensity H in the same way as electric induction or displacement is to electric intensity, it follows that the intensity at a point at a distance r from a pole of strength m in a medium whose permeability is μ, is given by

$$H = \frac{m}{\mu r^2}.$$

The magnetic induction is given by

$$B = \mu H = \frac{m}{r^2}.$$

Gauss's Theorem.—This states that the flux of magnetic induction across a closed surface is 4π times the total quantity of magnetism enclosed in that surface,

i.e. $\int B_n . dS = 4\pi \Sigma m$. [$B_n$ is the normal component of the magnetic induction at the element of surface δS considered.]

If the surface encloses one or more complete magnets $\Sigma m = o$, so that it is only when the surface cuts a magnet that the flux of magnetic induction across the surface is different from zero.

Lines and Tubes of Magnetic Induction.—A ***line of induction*** in a magnetic field is such that the tangent to it at any point indicates the direction of the magnetic induction at that point. A ***tube of induction*** is a tubular surface bounded by lines of induction.

Lines of Magnetic Induction used Quantitatively.—In the study of the relation between an electric current and the magnetic effects associated with it, it is often convenient to use lines of induction quan-

titatively. They are then imagined to be drawn in a uniformly magnetized medium in such a way that the number crossing unit area at right angles to the field is equal to the numerical value of the magnetic induction at that point. If the magnetization is not uniform it is necessary to consider an element of area δS at right angles to the direction of B at the point considered. Then δN, the number of lines of induction crossing this area is expressed by

$$\delta N = B \cdot \delta S.$$

Number of Lines of Induction from a Unit Magnetic Pole.—Let a closed sphere of radius r be constructed with a single magnetic pole of strength m at its centre. Let N be the number of lines of magnetic induction originating from m and crossing the surface above. Then the flux of induction across this surface is $4\pi r^2 \cdot B$, where B is the magnetic induction at any point on the surface of the sphere. But $B = \frac{m}{r^2}$.

$$\therefore\ N = 4\pi r^2 B = 4\pi m.$$

If the pole is in air, the lines of force become identical with the lines of induction, and we say that the number of lines of force arising from a unit pole in air is 4π.

The above result has been obtained without reference to Gauss's theorem because of its fundamental importance. Readers who are acquainted with the theorem will see at once that the result is true in general, for the flux of induction across a closed surface surrounding the pole m is $4\pi m$.

Magnetic Potential.—The magnetic potential at a point in a magnetic field is defined as the work done per unit positive pole against the field in bringing up a small positive magnetic pole from infinity to the point, the magnetic potential at infinity being considered to be zero.

The magnetic potential at a point in air, and at distance r from a pole m, may be determined as follows. The work done per unit positive pole against the field when a small positive pole moves from a point at distance r to another at distance $(r + \delta r)$ is

$$-\frac{m}{r^2} \cdot \delta r.$$

Hence, V, the potential at the point in question, is given by

$$V = -\int_{\infty}^{r} \frac{m}{r^2} \cdot dr = \frac{m}{r}.$$

The above result is only true for a point in air. If the point lies in a medium of permeability μ, the potential is given by

$$V = \frac{m}{\mu r}.$$

If V and $(V + \delta V)$ are the potentials at points distances x and $(x + \delta x)$ from a common origin and the medium is air, then δV is the work done in carrying unit positive pole from the point at lower potential to that at the higher potential. This is equal to $-H \cdot \delta x$ where H is the magnetic intensity between the two points (it is assumed it be uniform over the element of distance considered), the negative sign occurring since H is directed from the point at higher potential to that at the lower potential. Hence

$$H = -\frac{\partial V}{\partial x}.$$

It should be noticed that this expression is independent of the inverse square law.

If the position of the point is expressed in terms of its polar co-ordinates (r, θ), then H_r, the magnetic field strength in the direction of r increasing is given by

$$H_r = -\frac{\partial V}{\partial r}.$$

In a direction at right angles to this the element of length traced out by a point at distance r from the origin when θ becomes $\theta + \delta\theta$, is $r \cdot \delta\theta$. Hence H_θ, the magnetic field strength in this direction, is expressed by

$$H_\theta = -\frac{1}{r}\cdot\frac{\partial V}{\partial \theta}.$$

The Magnetic Potential at a Point in Air due to a Small Magnet.—Let NS, Fig. 42·2, be the small magnet of pole strength m, and let P be a point in air whose polar co-ordinates with respect to O, the centre of the magnet, are (r, θ) Then the potential at P due to N is m/NP; due to S it is $-m/\text{SP}$. But $\text{NP} = r - l\cos\theta$, and $\text{SP} = r + l\cos\theta$, where $2l$ is the length of the magnet.

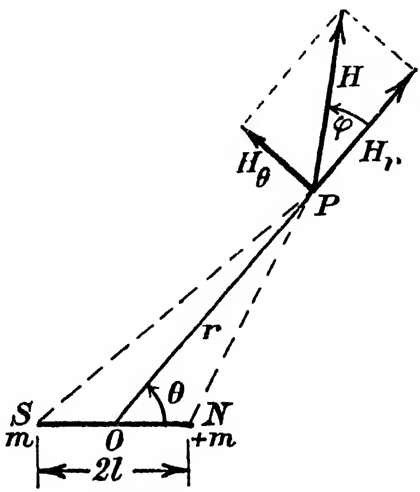


FIG. 42·2.—Magnetic Potential due to a small Magnet.

Hence $$V = m\left[\frac{1}{r - l\cos\theta} - \frac{1}{r + l\cos\theta}\right] = \frac{2ml\cos\theta}{r^2 - l^2\cos^2\theta}$$

$$= \frac{M\cos\theta}{r^2} \qquad \text{[if } l \text{ is small]}$$

where M is the magnetic moment of the small magnet.

If H_r and H_θ are the components of the magnetic field strength at P in the directions indicated, we have

$$H_r = -\frac{\partial V}{\partial r} = \frac{2M\cos\theta}{r^3},$$

and $$H_\theta = -\frac{1}{r}\frac{\partial V}{\partial\theta} = \frac{M\sin\theta}{r^3}.$$

The resultant field strength is therefore

$$\sqrt{H_r^2 + H_\theta^2} = \frac{M}{r^3}[4\cos^2\theta + \sin^2\theta]^{\frac{1}{2}} = \frac{M}{r^3}[1 + 3\cos^2\theta]^{\frac{1}{2}}.$$

If this makes an angle ϕ with H_r, $\tan\varphi = \frac{H_\theta}{H_r} = \frac{1}{2}\tan\theta$.

[Note that the resultant magnetic intensity is inclined to the initia line at an angle $(\phi + \theta)$.]

The Angle of Dip at a Point on the Surface of a Sphere when there is a Small Magnet at its Centre.—This is an important problem since the earth's magnetic field may, as a first approximation, be regarded as due to a small magnet at its centre. We shall therefore suppose that the negative pole of the small magnet points to the geographic north.—Cf. Fig. 42·3. Let P be a point on the surface in latitude λ (south). Then $(\theta + \lambda) = \frac{\pi}{2}.$ Hence the vertical component of the magnetic field at P is

$$H_r = \frac{2M}{r^3}\cos\theta = \frac{2M}{r^3}\sin\lambda.$$

The horizontal component of the magnetic field at P is H_θ, where

$$H_\theta = \frac{M\sin\theta}{r^3} = \frac{M}{r^3}\cos\lambda.$$

If ϕ is the angle which the resultant magnetic intensity, H_R, at a

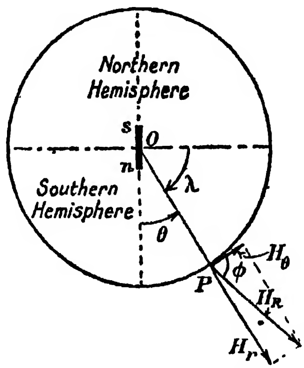


FIG. 42·3.—Calculation of the Dip in a given latitude (ideal case).

station in the 'southern hemisphere' makes with H_θ [ϕ is the angle of dip], then

$$\tan\phi = \frac{H_r}{H_\theta} = 2\cot\theta = 2\tan\lambda.$$

[In the 'northern hemisphere,' $\theta = \left(\frac{\pi}{2} + \lambda\right)$; the vertical component is negative, i.e. it is directed towards O. Also, $\tan\phi = -2\tan\lambda$, but, by convention, $0 < \phi < \frac{\pi}{2}$, so that the minus sign is neglected.]

EXAMPLE XLII

If ϕ_1 and ϕ_2 are the angles of dip observed in two vertical planes at right angles to each other and ϕ is the true dip, prove that

$$\cot^2\phi = \cot^2\phi_1 + \cot^2\phi_2.$$

CHAPTER XLIII

ELECTRICITY IN STEADY MOTION. CHARACTERISTIC PROPERTIES OF ELECTRIC CURRENTS. VOLTAIC CELLS

Electricity in Steady Motion.—In our study of electrostatic phenomena only electric fields which were practically invariable with respect to time have been contemplated. It is now necessary to investigate any effects which might be associated with the disappearance or annihilation of an electric field. Suppose that an electric field is due to a certain charged body: if this charge is removed the field becomes zero everywhere. Electricity has moved, i.e. an electric current has existed. The following experiment shows that an electric field does not always disappear at the same rate.

G, Fig. 43·1, is a gold-leaf electroscope arranged in parallel with a condenser [a Leyden jar for example]. The insulated plate of

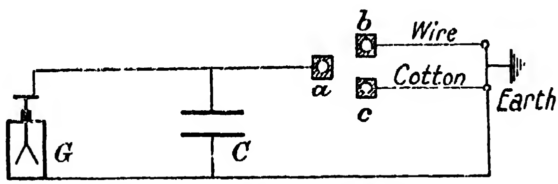


FIG. 43·1.—Electricity in Motion.

the condenser is connected to a small cavity *a* in a block of paraffin wax: the cavity contains mercury. Two similar cups, *b* and *c*, are connected to earth through a fine wire and a piece of cotton respectively. Suppose that *a* and *b* are connected by a copper wire—this must be supported on a sealing-wax handle to prevent the discharge of the condenser through the experimenter. The leaves of the gold-leaf electroscope collapse at once showing that the potential of the upper plate of C has been reduced to zero very quickly. If C is recharged, and *a* and *c* are connected by the copper wire, the collapse of the leaves takes place more slowly. In each instance we have the disappearance of a quantity of elec-

tricity and the electric field round it, but the rate of disappearance varies with the nature of the material along which the charge has been conducted to earth [or from earth to the condenser if the upper plate of the latter is negatively charged].

The Detection of Electric Currents.—Hitherto the presence of an electric current has been inferred from the disappearance of an electric charge: no direct means of establishing its existence has been mentioned. Let us now enumerate some means of detecting the presence of an electric current.

(i) ***The heating effect of a current:*** Suppose that a short length of very fine wire is stretched between two spheres, one insulated and the other earthed. If the knob of a charged Leyden jar is connected to the insulated sphere so that it is discharged the wire is volatilized with explosive violence.

If an experimenter, insulated by standing on blocks of paraffin wax, holds one knob of a Wimshust machine in action, the gas from a bunsen burner may be ignited if a copper wire held in the other hand is brought near to the escaping gas.

(ii) ***Mechanical effects:*** A sheet of glass or a piece of cardboard may be punctured when placed between the knobs of a Wimshurst machine in operation. The edges of the perforation in the cardboard will be burred outwards on both sides: the current is therefore oscillatory, i.e. there is a to-and-fro motion of the electric charges.

(iii) ***Chemical effects:*** Suppose a piece of filter paper, soaked in an aqueous solution of starch and potassium iodide, is supported on a piece of wax, and two wires, touching the paper, lead to the knobs of an electrical machine in action. Iodine is liberated when the discharge passes —this is indicated by the appearance of two blue patches at the points where the wires touch the paper. [From what occurs in the sequel it will be seen that the existence of two blue patches again indicates that the discharge is oscillatory.]

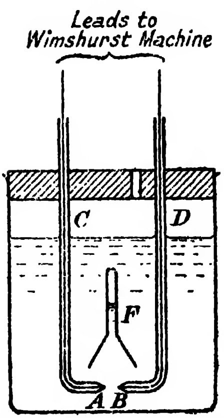


FIG. 43·2.—Chemical Effect due to an Electric Discharge.

Acidulated water [dilute sulphuric acid] may be decomposed by the passage of the discharge from a Wimshurst machine. Let A and B, Fig. 43·2, be the ends of two very fine platinum wires sealed into glass tubes C and D so that only the tips of the wires are exposed to the dilute acid in which they are immersed. If the discharge from an electric machine is passed across the gap AB for a long time bubbles of gas collect in F, a small funnel, the

delivery end of which is closed and very narrow. The solution has been decomposed—later, it will be learned that it is the water which has been decomposed, the amount of acid remaining constant.

(iv) ***Luminous effects:*** Suppose that Fig. 43·3 represents a glass tube containing air at a pressure of about 3 cm. of mercury. It is provided with electrodes A and B, these being platinum wires

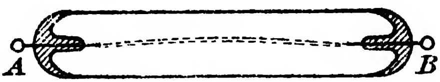


FIG. 43·3.—Luminous and Magnetic Effects of the Discharge.

sealed into the glass. If these are connected to a Wimshurst machine in operation, a long sinuous ribbon of light will be seen stretching almost along the complete length of the tube.

(v) ***Magnetic effects:*** If a cobalt steel magnet is placed near to the ribbon of light in the above tube, the path of the light will be distorted. Now it is a well-known scientific fact that only like things are affected by like things, i.e., in this instance, the passage of the electric current through the gas is accompanied by a magnetic field which is disturbed when a magnet is brought near to it.

ROWLAND, a physicist of the last century, found that when a series of insulated metal strips, mounted on a disc capable of revolution about its axis, are charged and the disc spun round the axis, a neighbouring magnetic needle is deflected. Such a motion of definite electric charges constitutes what is called an electric current, and the deflexion of the magnetic needle shows that moving electricity can be detected magnetically.

The Simple Voltaic Cell.—Suppose that a piece of amalgamated [1] zinc and a sheet of copper are dipped into dilute sulphuric acid. No action occurs—ordinary commercial zinc would dissolve owing to the 'local action' mentioned below. If the two plates are connected metallically the zinc begins to dissolve and bubbles of hydrogen appear on the copper plate. If the wire used to join the plates is thin it becomes hot; a small compass needle placed near the wire is deflected.[2] From these facts we conclude that there is a current flowing in the wire. The question presenting itself at once is this: Whence comes the energy to produce this motion of electricity? A condensing electroscope may be used to show that there is a difference of potential between the copper and zinc plates, the copper being positive. It is maintained by an *electromotive force,* e.m.f., in the cell. Later on [cf. Chap.

[1] An amalgam is defined as a solution of one or more metals in mercury.

[2] The effect can be increased by curling the wire so that it forms a spiral—or solenoid as it is termed.

XLVI] this will be further discussed : for the present it is sufficient to note that this electromotive force is measured by the potential difference between the copper and zinc when there is no current flowing.

Such a combination as

zinc | acid | copper

is known as a ***voltaic cell.*** The plates are called ***poles,*** or ***electrodes.***

It has been indicated above that the zinc strip dissolves as the current flows. The associated chemical reactions supply the energy necessary for the electricity to be sent through the wire and the cell itself. The current must flow through the acid as well as through the wire, otherwise there would be an accumulation of electricity at one or both of the electrodes which is contrary to experience.

The current is carried through the acid (the ***electrolyte***) by ***ions,*** i.e. atoms or groups of atoms which are charged. For example, sulphuric acid in water splits up into two hydrogen ions H^+, each carrying one elementary positive charge (equal and opposite to the elementary negative charge, termed an electron), and a sulphate ion carrying two negative charges, SO_4^{--}. Similarly, when common salt is dissolved in water, it ionizes :—

$$NaCl \rightarrow Na^+ + Cl^-.$$

In general, solutions of acids, bases, and salts, contain ions. We shall return to this question again in connexion with electrolysis.

Electrode Potentials—Nernst's Theory of Electrolytic Solution Pressure.—It has just been shown that a potential difference exists between two metals dipping into a dilute solution of sulphuric acid. Nernst first suggested (1889) a theory which would account for this and similar phenomena. According to this view every metal has a tendency to ionize, i.e. to acquire a charge of positive electricity when it passes into solution. Such an ion is termed a ***cation.*** With the noble metals this tendency is slight ; with copper it is greater ; with zinc still greater ; while with the alkali metals it is very high indeed. To each metal Nernst ascribed a definite ***electrolytic solution pressure*** which was a measure of the tendency which a particular metal had to form ions when placed in contact with water or an aqueous solution. This pressure is high for the alkali metals, low for gold, platinum, etc. For example, if a piece of zinc is dipped into pure water, a number of zinc ions (Zn^+) pass into the water in virtue of the fact that the zinc has an electrolytic solution pressure. In this process the loss of positive charge by the zinc rod causes the latter to become negatively charged, so that a definite potential difference exists

between the zinc and the water. Since the ions are charged particles they do not move away from the oppositely charged metal but form an electrical '*double layer*.' A state of equilibrium, in which the positive ions, being prevented from leaving the zinc owing to the negative charge acquired by the latter, form positive layers on the zinc, is very rapidly attained, and although the solution pressure is high, the mass of metal dissolved is too small to be detected by analytical methods. The equilibrium is a dynamic one.

Now consider what happens when a metal rod is placed in a solution containing its ions. Each square centimetre of the rod is bombarded by ions in solution, the number, n, striking per second depending upon the concentration of the ions and the temperature, i.e. upon the osmotic pressure of the solution. Each metal ion is positively charged, and if it adheres will give up its charge to the rod. On the other hand, depending upon the solution pressure of the metal, N metal ions will leave each square centimetre of the rod per second and pass into solution as positive ions, leaving an equivalent number of negative charges behind them. Three cases must be distinguished.

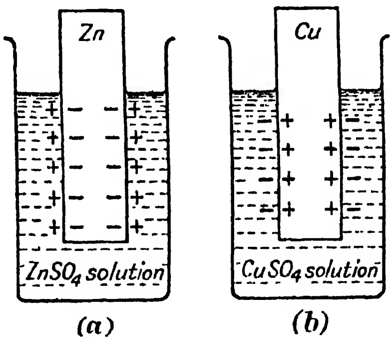


FIG. 43·4.

(i) If N is greater than n, there is a net transfer of ions from the metal into the solution, and this process goes on until the potential difference at the surface of separation between the metal and the electrolyte prevents the resultant transfer of ions. This state of affairs is very rapidly reached and then the electrode will have taken up an equilibrium negative potential with respect to the solution; this is the so-called ***electrode potential,*** whose value depends upon N, n, and the charge carried by each ion, as well as upon the temperature. Again, an electrical double layer has been formed and the mass of metal dissolving before equilibrium is reached is immeasurably small.

(ii) If $N = n$, the electrode potential will be zero.

(iii) If N is less than n, as will normally happen with copper and the noble metals, the electrode will acquire a positive potential with respect to the solution.

An example of (i) occurs when zinc is dipped into an aqueous solution of zinc sulphate, for example—cf. Fig. 43·4 (a). The lower the concentration of zinc ions in the solution, the lower will be

their osmotic pressure, i.e. the negative potential acquired by the rod will be increased.

When a copper rod is dipped into a solution of copper sulphate—Fig. 43·4 (*b*)—we have an example of (iii), and here the electrode potential will be greatest when the copper dips into a saturated solution of the copper sulphate, for then the osmotic pressure is a maximum.

Hitherto our remarks about solution pressure have been confined to metals, i.e. substances which produce positive ions. It is also applicable to substances which yield negative ions or ***anions.*** For example, chlorine yields negative ions, Cl^-, and if the solution pressure exceeds the osmotic pressure of the ions the electrode will acquire a positive potential. The solution pressure is, in such an instance, determined by the pressure of the gas forming on the electrode, so that the solution pressure is a function of the actual pressure of the gas—thus the electrode potential may be positive or negative, depending upon the pressure as well as upon the other factors mentioned above. It is very doubtful whether oxygen ions can exist as such in solution, but we can have OH^- ions formed :—

$$O_2 + 2H_2O + 4 \text{ electrons} \rightleftharpoons 4.OH^-.$$

The four electrons are carried over from the oxygen gas, leaving it positively charged. Oxygen in contact with water or an aqueous solution has a positive potential, whose value depends upon the pressure of the gas. Hydrogen forms H^+ ions. Hydrogen and oxygen electrodes can be used, as we shall see, the gas being supported by an inert solid such as platinum or carbon.

It is much easier to measure the p.d. between two electrodes than a single electrode potential, so Nernst proposed that all electrode potentials be referred to a standard hydrogen electrode in which hydrogen at a pressure of 76 cm. of mercury is in contact with a solution containing one gram-ion per litre. The table refers to some elements in equilibrium with solutions of 1 gram-ion per litre of the ion concerned.

Metal and Ions		Electrode Potential (Volt. with respect to Hydrogen)	Metal and Ions		Electrode Potential (Volt. with respect to Hydrogen)
Hg	Hg^{++}	+ 0·48	Ni	Ni^{+++}	− 0·22
O	OH^-	+ 0·42	Cd	Cd^{++}	− 0·40
Cu	Cu^{++}	+ 0·34	Fe	Fe^{++}	− 0·44
H	H^+	0	Zn	Zn^{++}	− 0·77
Pb	Pb^{++}	− 0·12	Na	Na^+	− 2·61

The Theory of the Simple Voltaic Cell and Polarization
-It can now be seen what happens when pure zinc and copper

are placed in the same jar of dilute acid. The zinc plate acquires a negative potential with respect to the solution, and the copper plate a positive one. Thus there is a p.d. between the plates equal to the e.m.f. of the cell. If the electrodes are connected by a wire, electrons will flow through it from the zinc to the copper and the equilibria at the electrodes will be upset. More Zn^{++} ions will pass into solution, and Cu^{++} ions will move towards the copper plate, there to lose their charges and be deposited. For every gram-ion dissolved or deposited, a definite quantity of electricity [cf. Chap. XLVII] will be transferred.

If the current taken from the cell is very small, the p.d. between the electrodes is maintained, but if the number of electrons reaching the copper plate from the wire is larger than can be neutralized by the deposition of Cu^{++} ions from the solution, the potential of this electrode becomes more negative, and the e.m.f. is lowered. The rate of deposition of Cu^{++} on the electrode depends on their concentration in the solution and upon the rate of diffusion from the bulk of the electrolyte to the electrode. Similarly, if the zinc is losing a large number of electrons per second, the concentration of the Zn^{++} ions near the electrode is increased and its electrode potential becomes less negative. Again the e.m.f. is lowered.

The lowering of the electrode potential of the copper may be sufficient for it to reach the electrode potential of hydrogen gas at atmospheric pressure. When this happens, hydrogen can exist in equilibrium with H^{+} ions, and some gas appears as bubbles on the copper plate. These bubbles act as 'insulators' on the electrode and make it still more difficult to take current from the cell. The positive electrode is now hydrogen, and the e.m.f. of the cell is low. When the e.m.f. is lowered on attempting to take a large current from the cell, the cell is said to be ***polarized.*** Polarization occurs in two stages, the first being due to a change in concentration of ions near the electrodes, and the second due to liberation of gas at the positive electrode. It should be emphasized that polarization is not *merely* the formation of hydrogen; the e.m.f. has been lowered considerably before that occurs. Polarization may be avoided by keeping the ion concentrations as constant as possible, and by oxidizing any hydrogen formed. The simple voltaic cell, copper and zinc plates in dilute sulphuric acid, polarizes very readily owing to the low concentration of copper ions.

The Daniell Cell.—The details of this cell, first described in 1836, are shown in Fig. 43·5 (*a*). The cell is very reliable and will supply a fairly steady current for a considerable time. The copper vessel, acting as the positive electrode, is in contact with a solution of copper sulphate, kept ***saturated*** by the copper sulphate crystals. The electrode potential of the copper is therefore kept constant.

The porous pot prevents the rapid mixing of the solutions. If dilute sulphuric acid is used, the electrode potential of the zinc is more negative and the e.m.f. of the cell is higher than if zinc sulphate solution is used, but in the latter case the e.m.f. is steadier, owing to the more constant concentration of the zinc ions. There is a very small p.d. across the junction of the two solutions. The e.m.f. of the cell is about 1·08 volts, depending on the temperature.

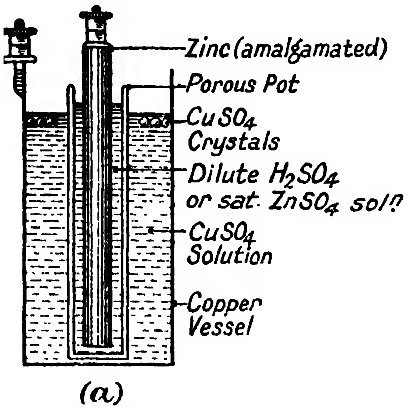


FIG. 43·5.—A Daniell Cell.

When the electrodes are put into metallic connexion with each other, electrons pass from the zinc to the copper, thus tending to make the potentials of the zinc and copper equal—cf. Fig. 43·5 (*b*). By convention, the current is considered to flow in the opposite direction. The current thus set up would cease after a small fraction of a second if further reactions did not occur. The electrons flowing away from the zinc cause the potential of this electrode to rise: the equilibrium at this electrode-electrolyte boundary is thereby disturbed and the changes occurring tend to maintain the electrode potential of the zinc, for zinc passes into solution. At the copper electrode the electrons arriving there make this electrode less positive, so that copper ions pass from the solution to it. Thus the concentrations of the solutions change when the cell is in action:

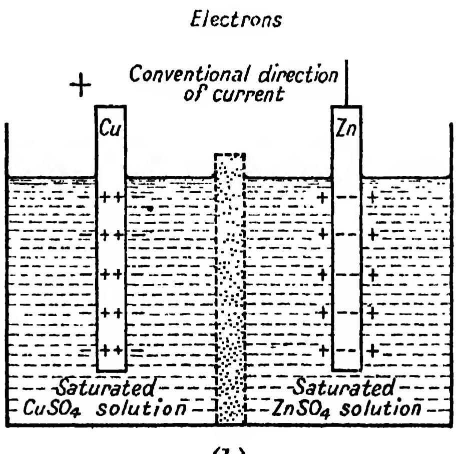


FIG. 43·5.—The Action of a Daniell Cell.

that of the zinc sulphate increases while that of the copper sulphate decreases. It is essential to keep the latter concentration as high as possible: hence the presence of the copper

sulphate crystals. The resultant chemical change in the cell is shown by the equation

$$Zn + CuSO_4 \rightarrow Cu + ZnSO_4.$$

If an inquiry regarding the source of electrical energy from the Daniell cell is begun it will be seen that the reaction

$$Zn + CuSO_4 \rightarrow Cu + ZnSO_4,$$

which represents the net chemical change when the cell is being discharged, is a reaction which occurs directly if zinc is actually put into a copper sulphate solution. It is, in fact, an ordinary spontaneous chemical process which takes place readily if the reacting substances are placed in contact: in this instance, of course, no electrical energy is produced but merely heat—the ordinary ***heat of reaction***. In the Daniell cell this same chemical action is made to occur in an abnormal fashion. The reacting substances are not in direct contact, but there is an electrically conducting path between them. When the circuit is completed the reaction takes place in two separate parts—dissolution of zinc from one pole, deposition of copper on the other. The net chemical result is the same as in the direct reaction, but in this instance electrical energy is obtained instead of heat.

In either case this energy comes from the chemical energy residing in the system, $Zn + CuSO_4$, and it is this energy which brings about the chemical change. In the direct reaction this appears as heat, in the cell reaction as electrical energy. Similarly any voltaic cell may be regarded as a device whereby some spontaneously occurring chemical reaction is ***harnessed*** so that its chemical energy appears as electrical energy instead of as heat.

The Leclanché Cell.—A cell of higher e.m.f. than a Daniell cell must have electrodes with a greater difference of potential than there is between copper and zinc in dilute sulphuric acid. Metals having a more negative electrode potential than zinc dissolve too quickly to be of much service. Oxygen gas, however, has a higher electrode potential than copper, and there is a number of cells using an oxygen positive electrode, of which the Leclanché is the most important. A diagrammatic section of this cell, whose e.m.f. is about 1·46 volts, is shown in Fig. 43·6. The oxygen is supported on a carbon rod. Carbon occludes and 'adsorbs' oxygen very readily, but cannot take part in the cell reactions because it does not ionize. In the wet form of the cell, the carbon rod is in a porous pot containing a mixture of manganese dioxide, MnO_2, and powdered graphite, which is necessary to conduct the current, manganese dioxide being a poorly conducting substance. The porous pot and zinc rod stand in a saturated solution of sal ammoniac, NH_4Cl, in a glass jar, the upper parts of which are coated with a

suitable enamel to prevent creeping'. The chemical reactions may be summarized as

$$Zn + 2NH_4Cl + 2MnO_2 \rightarrow Zn(NH_3)_2Cl_2 + H_2O + Mn_2O_3.$$

The MnO_2 is thus reduced to Mn_2O_3, which is slowly oxidized back again by contact with the atmosphere. If the MnO_2 is reduced too rapidly, by taking a large current from the cell, the

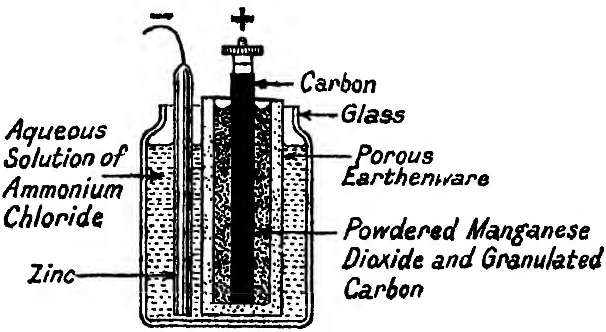


FIG. 43·6.—A Leclanché Cell.

oxygen may be taken from the carbon electrode instead of from the air, and the electrode potential falls (possibly so far that hydrogen is liberated, and the electrode becomes a hydrogen one. The manganese dioxide also helps to keep the concentration of the hydrogen low). The cell, therefore, can be used only for intermittent currents, but it will stand for long periods without attention.

The Dry Cell.—This cell is really a form of Leclanché cell in which the fluid has been replaced by a mixture of sal-ammoniac, hygroscopic salts, and sawdust. This mixture must be *moist*, so that the term 'dry cell' is really a misnomer; its very action depends upon the fact that it must be wet. The carbon rod is surrounded by a paste made from manganese dioxide, coke, ammonium chloride and zinc chloride, this being a hygroscopic substance. This depolarizing paste is contained in a muslin bag (for the paste is a good conductor and must not be allowed to come into contact with the zinc). The mixture of sal-ammoniac, zinc chloride, and sawdust, occupies the small space between the bag and the outer zinc case which forms the negative electrode. A small vent in the wax which seals the cell permits any gases to escape.

In the making of a dry cell one of the main considerations is to ensure the retention of the essential moisture in the interior of the cell, while at the same time permit the gases generated during the working of the cell to escape. In the cell shown in Fig. 43·7 these conditions are adequately fulfilled. There is a patented device which hermetically seals in the active ingredients but which permits the gases to escape during the working of the cell. In addition it is possible to introduce a large quantity of the depolarizing mixture by

dispensing with the usual fabric sack and separating the mixture from the zinc container by a thin paper lining. The whole is moistened with an aqueous solution of ammonium chloride and when the paper is saturated with this solution the current passes through the paper—it is then a separator but not an insulator.

The central carbon rod has two narrow holes running longitudinally through it. At the bottom of the cell a cardboard disc fits closely round the rod and presses against the paper lining. The depolarizing mixture is then rammed into the space between the rod and paper lining. A layer of the sealing plastic compound is then run on top of the mixture. Gases formed in the mixture rise until they reach the plastic layer and when the pressure becomes great enough they pass as bubbles through the compound. These burst, and the compound flows together, thereby re-sealing the cell.

The air space at the top of the cell connects with the vents through the carbon by means of slots cut in the rod. The reason for the cardboard disc with grooves at the bottom of the cell is that by this means any gases generated at the base of the cell find an exit through the vents in the carbon. The depolarizing agent is packed in so firmly that the gases formed in this region are unable to pass through the mixture, and, unless they can escape in another way, gradually force the mixture out of the cell.

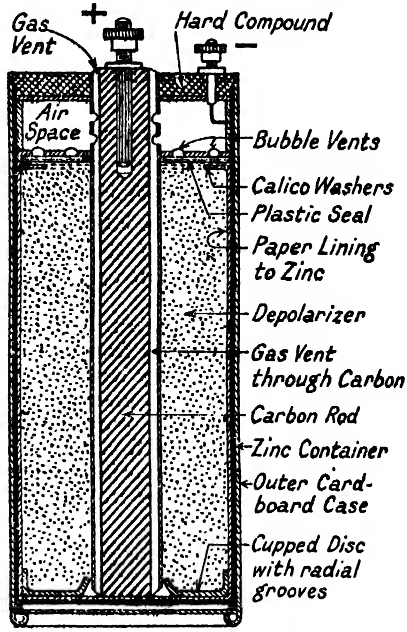


FIG. 43·7.—Siemen's Dry Cell.

[Another advantage of this new type of dry cell is that its internal resistance is low—on short circuit a current of 50 amperes is obtained.]

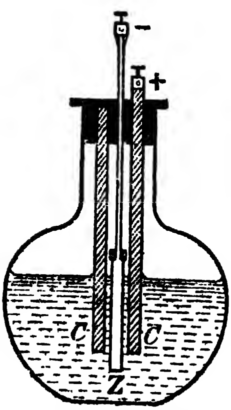


FIG. 43·8.—A Bichromate Cell.

The Bichromate Cell.—In this cell, Fig. 43·8, carbon and zinc are employed as the two electrodes. The liquid is dilute sulphuric acid in which potassium bichromate has been dissolved. The solution may be made as follows: 1000 cm³. water, 100 cm³. conc. H_2SO_4, 80 gm. $K_2Cr_2O_7$. The bichromate acts as the depolarizing agent, the CrO_3 constituent being reduced

to Cr_2O_3. The chemical action of the cell is represented by

$$3Zn + K_2Cr_2O_7 + 7H_2SO_4 \rightarrow 3ZnSO_4 + K_2SO_4 + Cr_2(SO_4)_3 + 7H_2O.$$

The e.m.f. of this cell is approximately 2 volts.

The Bunsen Cell.—A porous pot containing concentrated nitric acid and a carbon electrode [+] is immersed in a vessel containing dilute sulphuric acid. A zinc rod is placed in the sulphuric acid and forms the negative electrode of the cell. The nitric acid is reduced by the hydrogen which is formed when the zinc dissolves, and is therefore an efficient depolarizing agent. These cells are very objectionable in a laboratory on account of the nitrous fumes which are evolved. If the carbon rod is replaced by a sheet of platinum, then the cell is as designed by GROVE.

The Weston Cadmium Cell.—The cells which have been described previously suffer from the disadvantage that their e.m.fs. are not constant when they are in use, and also vary considerably with changes in temperature and changes in the concentration of the dissolved substances. For purposes of standardization it is desirable to have a cell whose e.m.f. shall be constant, or, if it does vary with temperature, then this variation must be small and measurable. Such a constant cell is found in the Weston Cadmium Cell. A cadmium

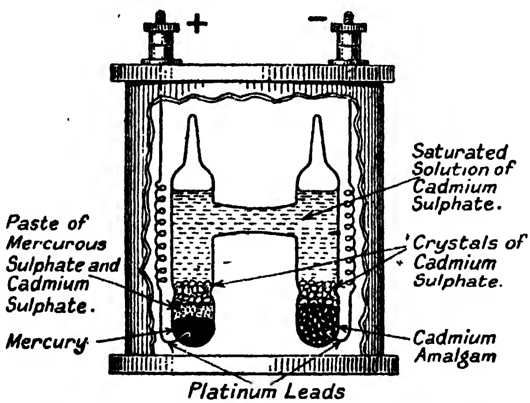


FIG. 43·9—Weston Cadmium Standard Cell.

amalgam forms the negative pole, whilst mercury is the positive pole. The liquid in the cell is a saturated solution of cadmium sulphate, and mercurous sulphate is the depolarizer. All these substances are specially purified before being assembled as in Fig. 43·9. Platinum wires serve to connect the electrodes to an external circuit. Such cells have an exceptionally high internal resistance. They are not intended to give any but very minute currents, and, in spite of their high internal resistance [cf. Chap. XLVI], are spoiled if the terminals are connected by a short wire. They are only used as standards with

which other cells may be compared. The e.m.f. of such a cell is expressed by

$$E_t = 1{\cdot}0186 - [3{\cdot}8 \times 10^{-5}(t - 20)] \text{ volt.},$$

where E_t is the e.m.f. at $t°$ C. From the formula it is seen that the e.m.f. is 1·0186 volts at 20° C.

In all forms of cells in which zinc is used as one plate, or ***electrode***, the zinc gradually dissolves, unless it is exceptionally pure, even when the two electrodes are not connected together. The high cost of production of very pure zinc renders its use prohibitive ; it has been found, however, that if the zinc contains 4 per cent. of mercury then the zinc only dissolves when the cell is in use. The solution of commercial zinc in the sulphuric acid, when the cell is not operating, is referred to as ***local action***, which we may explain as follows.—Commercial zinc contains traces of iron and other metals. If such an impurity is on the surface of the zinc plate, and therefore in contact with the acid, it will behave as the positive electrode of a small voltaic cell. In this small cell the zinc will be the negative electrode and will be dissolved even when the copper and zinc plates of the large cell are not connected together.

Whenever a steel framework is exposed to the action of water containing traces of dissolved salts the metal is gradually corroded away. Attempts have been made to prevent this by connecting the iron structure with a mass of zinc likewise exposed to the same water. It was thought that the zinc alone would corrode and thus save the iron structure. Experience has shown that the zinc is only effective for the iron in its immediate neighbourhood.

EXAMPLES XLIII

1.—Describe a condensing electroscope and explain its use.

2.—Describe the simple voltaic cell and give an account of its action. Explain how, and to what extent, the defects of the simple cell are remedied in (*a*) a Daniell cell, (*b*) a Leclanché cell.

CHAPTER XLIV

ELECTRIC CURRENTS AND THEIR MAGNETIC EFFECTS

The Magnetic Field due to a Current in a Straight Wire.—When an electric current flows in a straight wire magnetic forces are produced in the neighbourhood of the wire. The wire itself does not become a magnet, for it cannot attract iron filings, neither does it possess any magnetic poles. If, however, a vertical wire carrying a *large* current pierces a sheet of cardboard on which iron filings have been sprinkled, then these filings arrange themselves in circles round the wire. The filings are still arranged in a circular form when the current is reversed; but if a small compass needle [which may be regarded as an iron filing capable of rotation about a pivot] is placed on the cardboard, the direction in which the needle points depends upon the direction of the electric current in the wire. The direction of these circular magnetic lines of forces can be ascertained from the following rule: ***Look along the wire in the direction in which the electric current is travelling, then the lines of force are such that a positive (north-seeking) pole tends to move in a clockwise direction.*** In Fig. 44·1 the direction in which the N-pole of a needle tends to move is indicated. The manner in which the magnetic field is related to the direction of the current is perhaps best remembered with the aid of Fig. 44·2. [If, on this diagram, the names 'current' and 'magnetic field' are interchanged, we have the direction of the magnetic field at the centre of a circular coil carrying a current flowing in the direction indicated.]

The tendency of a magnetic pole to rotate may be demonstrated by means of the following experiment [Fig. 44·3]. A straight wire W dips into the centre of a vessel containing acidulated water and is connected to the one pole of a battery. [N.B. A cell is denoted by two parallel lines, one shorter and thicker than the other. The longer line represents, by convention, the positive electrode.] A copper plate P is joined to the second electrode. A piece of soft iron rod is magnetized and inserted in a cork C, the shape of the cork being such that the rod floats vertically, with one pole near the sur-

face and the other well below it. When a strong current is passed through the wire the floating magnet moves in a circle as long as the current is passing. The direction of the motion depends upon the polarity of M and the direction of the electric current in W.

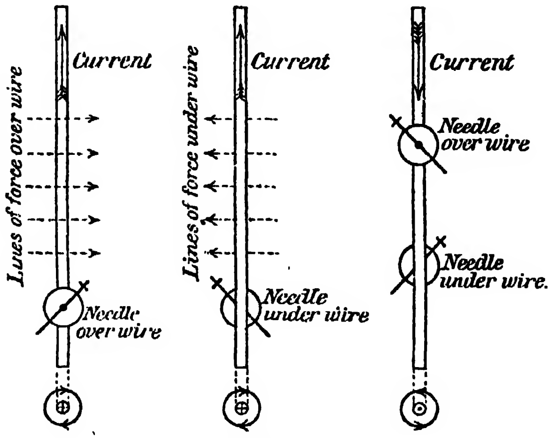


FIG. 44·1.

[N.B.—The cross in the small circle represents the end of an arrow which is moving away from the observer; the dot in its circle represents the point of an arrow when this is seen approaching.]

In the above experiments a large current is used so that the effect of the earth's magnetic field may be small compared with that due to the current. The same effect may be obtained by passing a somewhat weaker current through several wires in parallel.

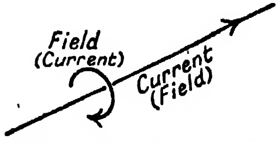


FIG. 44·2.—A steady Current and its Magnetic Field.

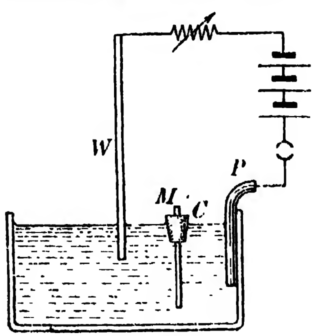


FIG. 44·3.

If the current is weak and only one wire is used the magnetic field in a horizontal plane round the wire may be plotted with the aid of a compass needle. Such a field is indicated in Fig. 44·11. The presence of a neutral point will be noticed.

Maxwell's Rule.—Maxwell gave the following rule for determining the direction of the magnetic field due to a current flowing in a specified direction :—***If an observer imagines that a corkscrew is being driven in the direction of the current, a north pole, placed in the field, will move in the same direction as the screw is being turned.***

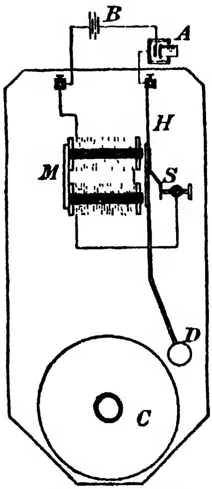


FIG. 44·4.—Electric Bell Circuit.

Electric Bells.—The electric bell is a simple practical application of the magnetic effect of an electric current : the construction and mode of action of such a bell are as follows :—In Fig. 44·4 M is an electromagnet excited by the current from a battery B when the button of the switch A is pressed inwards so that contact is made between two small metal plates in it. The current flows through a spring H which is normally in contact with an adjustable contact S, and then through the coils of the electromagnet back to the battery. Attached to the spring H there is a piece of soft iron which is attracted to M when the current is established. If the contact between H and S has been properly adjusted this contact is broken when the soft iron moves towards M and the current ceases. The magnet is no longer excited, H moves back to its normal position, and the whole process is repeated. Attached to H is a hammer D which strikes the gong C and continues to vibrate until the pressure on A is released.

Telegraphy.—Fig. 44·5 shows in simple form the equipment at two stations between which signals have to be sent. It will be noticed that the equipment is the same at each station and that the stations are connected by a wire or 'line.' Until 1837 a return wire was used to enable the current to return to the sending station, but in that year STEINHEIL (Munich) discovered that the earth was sufficiently conducting to be used for that purpose. The apparatus consists of a Morse key ACD movable about a horizontal axis through C. Normally there is contact between A and a metal stud F, but this contact is destroyed when D is depressed to make contact with H. M is an electromagnet and the current is supplied from a battery B. G is a galvanometer to indicate to the observer that a current is passing along the line L. When contact is made between D and H the electromagnet at the sending station is not in action, but the electromagnet M′ is excited. This pulls downwards a piece of iron S′—the armature. S′ is attached to a lever carrying an inked wheel W′. When M′ is excited, i.e. signals are coming in from the sending station, a narrow piece of paper is made to move automatically below

the wheel W′. The speed at which this moves is regulated so that if D is pressed down for a short time a dot is registered. If D is held down for a longer time the wheel makes a dash. A prearranged code of dots and dashes enables a message to be sent.

If signals were sent over a long distance with the aid of an apparatus similar to that just described they would be so feeble when they

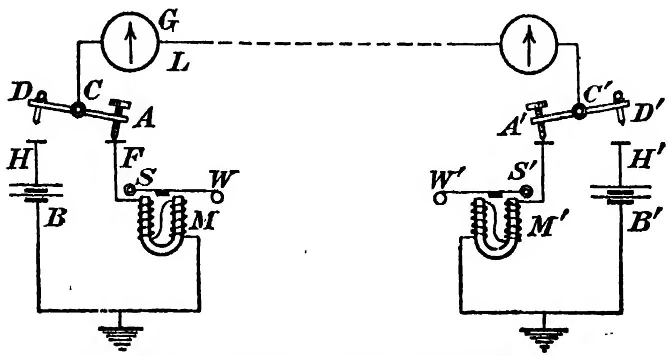


FIG. 44·5.—Telegraph Circuit with Morse Key.

arrived at the second station that the instruments there would not respond to them. This necessitates the use at the receiving station of a relay which consists of an electromagnet, the armature being delicately adjusted. The feeble signals cause this armature to move and close a local circuit in which the current is so strong that the receiving Morse instrument indicates the arrival of the signals.

Magnetic Shells of Uniform Strength.—When a thin sheet of magnetizable substance of uniform thickness is magnetized in a direction perpendicular to the surface of the sheet we have what is termed a ***magnetic shell.*** It is now necessary for us to investigate the properties of such a shell since its magnetic effects are equivalent to those of a current of certain strength flowing in an electric circuit coinciding with the periphery of the shell. Although such shells do not actually exist they are a means of correlating the phenomena of electric currents and magnetism.

The strength, ϕ, of a uniformly magnetized shell is defined as its magnetic moment per unit area. Thus

$$\phi = \frac{\text{magnetic moment}}{\text{area}} = \frac{\text{intensity of magnetization (J)} \times \text{volume}}{\text{area}}$$

$= Jt$, where t is the thickness of the shell.

Magnetic Potential due to a Magnetic Shell at an External Point.—Just as the electric potential at a point in an electric field is defined as the work done against the field per unit positive charge in bringing up a small positive charge from infinity to that point, so is the magnetic potential at a point in a magnetic field

defined as the work done against the field per unit positive pole in bringing up from infinity a small positive pole to that point. Hence, by an argument similar to that given on p. 683, the magnetic potential due to a magnetic pole of strength m at distance r from it is $\frac{m}{r}$. Let P, Fig. 44·6 (*a*), be a point at a distance r from the centre of a small element ABCD of a uniform magnetic shell RS; let θ be the angle made by r with the axis of magnetization of the element and let δs be the area of each face of the element. Then the amount of magnetism on each face is numerically equal

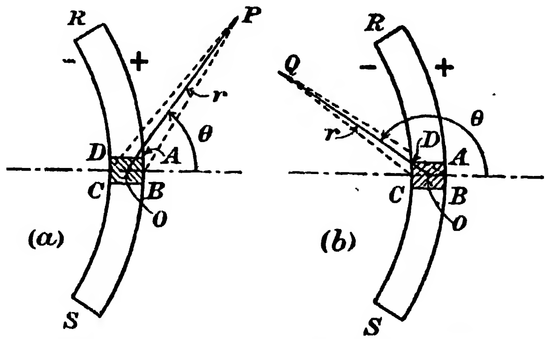


FIG. 44·6.—Magnetic Potential due to a Uniformly Magnetized Shell.

to J . $\delta s = \delta m$ [say]. The magnetic potential at P, a point on the positive side of the shell, due to the charge on AB, is $\frac{\delta m}{r - l\cos\theta}$; that due to the charge on CD is $-\frac{\delta m}{r + l\cos\theta}$, where $2l = t$, the thickness of the shell.

$$\therefore\ \delta V_P = \delta m\left[\frac{1}{r - l\cos\theta} - \frac{1}{r + l\cos\theta}\right],$$

where δV_P is the contribution to the magnetic potential at P, due to the charges considered. Since $l/r \to 0$, we have,

$$\delta V_P = \frac{\delta m \,.\, 2l\cos\theta}{r^2} = \frac{Jt\cos\theta \,.\, \delta s}{r^2}$$

$$= \frac{\phi\delta s \,.\, \cos\theta}{r^2}.$$

But $\delta s \,.\, \cos\theta/r^2 = \delta\omega$, where $\delta\omega$ is the solid angle subtended by the element of the shell at P. Hence

$$\delta V_P = \phi \,.\, \delta\omega.$$

Since this is true for each element of the shell, we have,

$$V_P = \phi\omega,$$

where V_P is the magnetic potential at P, and ω is the solid angle subtended by the shell at P.

Suppose now that we require the magnetic potential, V_Q, at Q, a point on the negative side of the shell. Then from Fig. 44·6 (*b*), as before,

$$\delta V_Q = \delta m\left[\frac{1}{r + l\cos(\pi - \theta)} - \frac{1}{r - l\cos(\pi - \theta)}\right]$$

$$= \phi \cdot \frac{\delta s \cdot \cos\theta}{r^2} \qquad \text{[as before]}$$

Now $\quad \delta\omega = \dfrac{\delta s\cos(\pi - \theta)}{r^2} = -\dfrac{\delta s \cdot \cos\theta}{r^2}$

$$\therefore\ \delta V_Q = -\phi \cdot \delta\omega,$$

and $\quad V_Q = -\phi \cdot \omega.$

The meaning of this equation is that if we approach the shell and bring up a small magnetic pole from an infinite distance and finally arrive at a point on the negative side of the shell, the work done against the field per unit positive pole is $-\phi\omega$.

Equivalent Magnetic Shells.—Let us assume that a current i is flowing in a circuit S, Fig. 44·7. Let us imagine that this has been replaced by the network shown and that a current i is flowing round each mesh such as the one shown shaded in the same direction as the original current. We are justified in this for each line separating two adjacent elements is traversed by equal currents in contrary directions so that any magnetic effect due to one is neutralized by the other, and it is only where the elements touch the periphery that the effects are not neutralized in this way. Hence the effect of all the elementary currents is the same as that of the current in the original circuit.

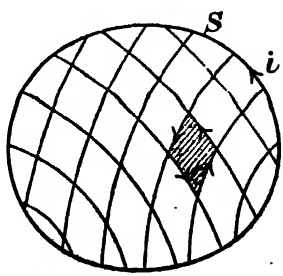


FIG. 44·7.—Equivalent Magnetic Shells.

Now experiment shows that at distances from such elements great compared with their linear dimensions the magnetic effect is the same as that of a suitably chosen magnetic shell whose boundary coincides with that of the element. Since every closed circuit may be conceived as made up of contiguous elements as above it follows that in as far as its magnetic effects are concerned every closed circuit may be replaced by an equivalent magnetic shell [provided that the point where the field is considered does not lie inside the shell].

The Electromagnetic Unit of Current.—When the magnitude of the current flowing in a closed circuit is numerically equal to the strength of its equivalent magnetic shell, that number represents the magnitude of the current in electromagnetic units. Hence *the unit e.m. current is that which, when flowing in a closed circuit, may be replaced by an equivalent magnetic shell of strength unity.*

Work done per Unit Pole in Carrying a Small Positive Magnetic Pole round a Closed Circuit.—Let ABC, Fig. 44·8, be a circuit carrying a current i e.m.u. This circuit may be replaced by any uniform magnetic shell $[\phi = i]$ whose boundary coincides with that of the circuit, in so far as the magnetic field outside the region not occupied by the shell is concerned. Let ADB be such a shell. Consider two points P_1 and P_2 very close to the magnetic shell, but on opposite sides of it. We shall assume that P_1 is on the positive side of the shell. Then the magnetic potential at P_1 is $\phi\omega_1$, where ϕ is the strength of the shell, and ω_1 the solid angle subtended by the boundary of the shell at P_1. At P_2 the potential is $-\phi\omega_2$. Hence the work done per unit positive pole in taking a small positive pole from P_2 to P_1 by a path not cutting the shell is

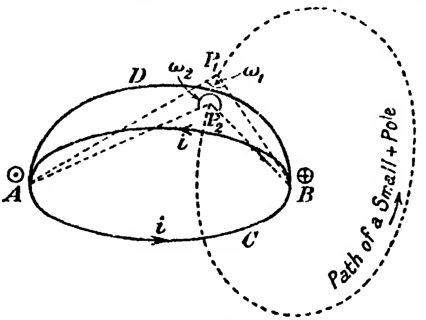


FIG. 44·8.

$$\phi\omega_1 - (-\phi\omega_2) = \phi(\omega_1 + \omega_2).$$

In the limit when the shell is infinitely thin, its strength remaining equal to ϕ, however, the points P_1 and P_2 practically coincide,

$$\therefore\ \omega_1 + \omega_2 = 4\pi.$$

Hence the work done is $4\pi\phi$.

[It should be noted that if the magnetic pole is moved from P_1 to P_2 by a path not intersecting the shell, the work is done by the field.]

If we now pass from P_1 to P_2 through the shell, the direction of the field is reversed and the work done per pole unit is $-4\pi\phi$. It is not necessary to establish this analytically, for the magnetic potential at a point due to a shell must be single-valued, so that the total work done when the path is completely closed is zero. Otherwise useful work could be obtained by allowing a magnetic pole to move round a closed path drawn in a magnetic field ; this is contrary to experience.

When we are dealing with the work done in threading a closed circuit, however, there is no actual shell present, and the work done in passing from P_1 to P_2 to complete the path is zero. Under these circumstances, the work done on a unit positive magnetic pole in passing from P_2 to P_2 via P_1, i.e. in completely threading the circuit, is $4\pi\phi = 4\pi i$.

In this instance the magnetic potential at a point will be a multi-valued function, i.e. its potential may be any of a series whose values

differ by multiples of $4\pi i$. This is not contrary to the principle of the conservation of energy, for a current is not a static effect, and the circuit is not in the same condition as it was initially, when it has been threaded by a magnetic pole.

The Magnetic Field due to a Circular Current at a Point on its Axis.—Let P, Fig. 44·9 (*a*), be the point considered. Let i be the current in electromagnetic units, and r the radius of the circle. The magnetic potential at P, a point on the positive side of the equivalent magnetic shell, is therefore $i\omega$, where ω is the solid angle subtended at P by the boundary of the circle. Now the solid angle ω is formed by the revolution of the diagram about

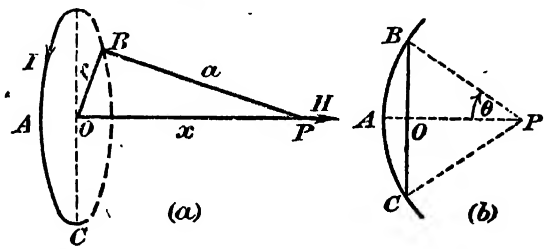


FIG. 44·9.—Field due to a Circular Current.

the axis OP. It is defined as the area of the spherical cap ABC, Fig. 44·9 (*b*), divided by the square of the distance BP. Let $BP = a$. Hence

$$\omega = \frac{2\pi a^2(1 - \cos\theta)}{a^2} = 2\pi(1 - \cos\theta) = 2\pi\left[1 - \frac{x}{(r^2 + x^2)^{\frac{1}{2}}}\right].$$

The magnetic intensity at P is $-\dfrac{dV}{dx}$, or $\dfrac{2\pi i r^2}{(r^2 + x^2)^{\frac{3}{2}}}$. The intensity, $H_{x=0}$, at O is found by putting $x = 0$ in the above. We have $H_{x=0} = \dfrac{2\pi i}{r}$. If there were N turns of wire, the magnetic field at O would be increased N-fold, i.e. $H_{x=0} = \dfrac{2\pi N i}{r}$.

The Magnetic Field due to a Linear Current.—Let P, Fig. 44·10, be a point at distance r from an infinite wire carrying a current i. Let this current flow downwards. The return wire may be considered as B at infinity. The equivalent magnetic shell in this instance is one half of an infinite plane. If θ is the $\widehat{APB}$, the solid angle subtended at P by the plane is 2θ since 2θ is the area of a unit sphere whose centre is P cut off by planes PA and PB normal to the paper. The section of the sphere by these planes is like a section of an orange. Since the surface area of a hemisphere of unit

radius is 2π and π is the angle subtended at its centre by its curved surface, it follows that the solid angle is 2θ in the present instance

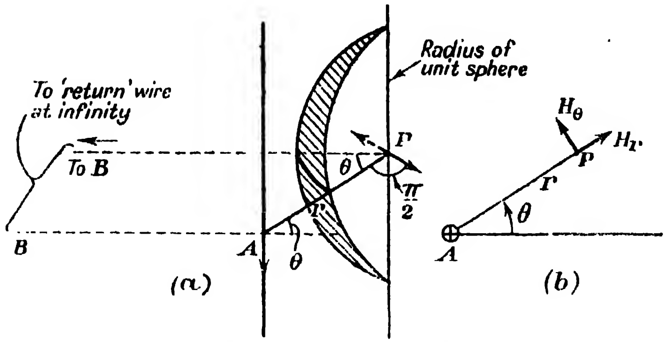


FIG. 44·10.—Magnetic Field due to a Linear Current.

since a hemisphere may be considered as a number of congruent sections side by side. The magnetic potential at P is $2i\theta = V$.

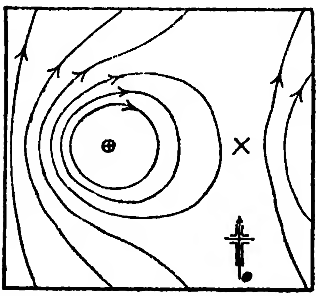


FIG. 44·11.—Combined Magnetic Field due to a Vertical Linear Current and H_0.

The magnetic field strength, H, at P is $-\frac{1}{r}\cdot\frac{dV}{d\theta}$, since $r.d\theta$ is the element of length at P in the direction of θ increasing. This is $-\frac{2i}{r}$ [indicated by dotted arrow at P]. Hence the lines of magnetic force are circles. The actual direction of this intensity is shown by the full-line arrow.

The above argument is only true when the earth's magnetic field is absent. When the earth's field is present the lines of force are distorted as in Fig. 44·11. If H_0 is the strength of the earth's horizontal field and r the distance of the neutral point[1] from the wire, we have $H_0 = \frac{2i}{r}$, so that i may be determined. Its value will be in electromagnetic units of current.

If a small compass needle is placed at a distance r due N (or S) of a vertical wire carrying a steady current, the needle will take up an equilibrium position determined by H_0 and H, two magnetic fields mutually perpendicular. If θ is the deflexion of the needle

[1] The point is readily and accurately determined by constructing two directional loci as explained on p. 756.

$H = H_0 \tan\theta$. If we assume $H = Ar^n$, where A and n are constants, we may write

$$\log H_0 + \log \tan\theta = \log A + n \log r.$$

By taking a series of observations and then plotting $y = \log \tan\theta$ against $x = \log r$, a straight line of slope n will be obtained. It will be found that $n = -1$, as the above theory requires.

A Modern Variation of Biot and Savart's Experiment.—The manner in which the magnetic field due to a long vertical straight wire carrying a current varies with the distance r from the wire may be examined experimentally with the aid of a vibration magnetometer. This was originally done by BIOT and SAVART. Suppose that A, Fig. 44·12, is a section of the wire, the current flowing towards the observer. Suppose that the centre of the magnetometer needle is directly east of the wire, and at a point P distance r from it. Let H be the field due to the current at this point. Then the total horizontal field at P is $H_0 + H$, if H_0 is the strength of the earth's horizontal magnetic field.

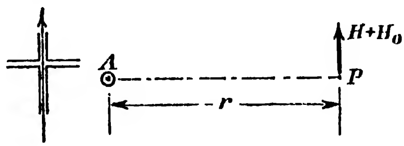


FIG. 44·12.—Biot and Savart's Experiment.

Suppose that the magnetic needle has a period T_0 when there is no current in the wire. Then

$$T_0 = 2\pi\sqrt{\frac{I}{MH_0}},$$

where M is the magnetic moment of the magnet and I its moment of inertia about the axis of rotation. When there is a current in the wire, let T be the period of oscillation. Then

$$T = 2\pi\sqrt{\frac{I}{M(H_0 + H)}}.$$

From these equations, we have

$$\left(\frac{T_0}{T}\right)^2 - 1 = \frac{H}{H_0}$$

Assume $H = Ar^n$, where A and n are constants. Then

$$\log\left[\left(\frac{T_0}{T}\right)^2 - 1\right] = C + n \log r,$$

where C is a constant.

If we plot $y = \log\left[\left(\frac{T_0}{T}\right)^2 - 1\right]$ and $x = \log r$, the slope of the straight line is n. It will be found equal to -1, i.e. the magnetic

field at a point due to the linear current varies inversely as the distance of that point from the conductor.

Maxwell's Experimental Proof.—Maxwell suggested that the following experiment might be used to show that the magnetic field at a point due to a long straight current was inversely proportional to the distance of the point from the wire. The apparatus consists of a cardboard ring A, Fig. 44·13, supported by three equal strings so that it is free to rotate in a horizontal plane. Suppose that B is the wire carrying the current. Let NS be a magnet placed on the cardboard so that it lies in the magnetic meridian with its north-seeking and south-seeking poles at distances r_1 and r_2 respectively from the wire. C is a non-magnetic counterpoise. Let H_1 and H_2 be the magnetic field strengths at distances r_1 and r_2 from the wire.

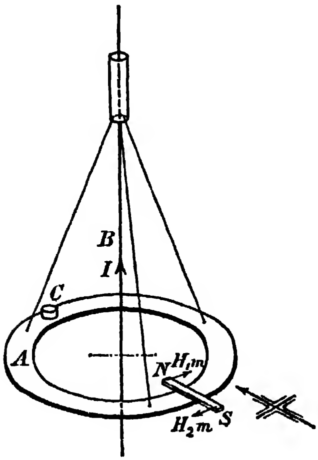


FIG. 44·13.—Maxwell's Apparatus for Investigating the Magnetic Field due to a Vertical Linear Current.

Then the forces on the magnet, due to the current in B, are H_1m and $-H_2m$. The moment of these forces about the axis of suspension is

$$H_1mr_1 - H_2mr_2.$$

Experiment shows that there is no tendency for the system to rotate. Hence

$$H_1r_1 - H_2r_2 = 0,$$

i.e.

$$\frac{H_1}{H_2} = \frac{r_2}{r_1},$$

or the magnetic field varies inversely as the distance from the wire.

[It is not necessary for the wire to pass through the centre of the ring—why ?]

Magnetic Field inside a Long Straight Solenoid wound Uniformly.—Consider such a solenoid having n turns per unit length. Let i be the current in e.m.u. Then each turn of the wire and the current in it may be replaced by its equivalent magnetic shell. Let us assume that these shells are normal to the axis of the solenoid, and that opposing faces of adjacent shells are very close together. Some such shells are shown in Fig. 44·14 (*a*). The strength of these equivalent shells is i, i.e. the magnetic moment

per unit area of a face is i. The magnetic moment of a shell is therefore $\pi a^2 i$. But the thickness of the shell is $\frac{1}{n}$.

$$\therefore \text{ pole strength} = \pm \pi a^2 i \div \frac{1}{n} = \pm n\pi a^2 i.$$

Let σ be the surface density of the distribution of magnetism on the faces of the equivalent shell. Then

$$\sigma = \pm ni.$$

Consider a point P near the centre of the solenoid and in the space between two adjacent shells. Since the distance between

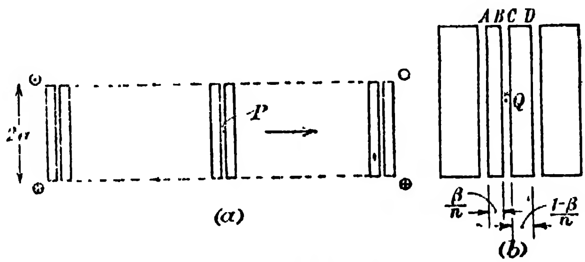


FIG. 44·14.—Magnetic Field inside a Long Straight Solenoid wound Uniformly.

these faces is small, the intensity, H, at P is $+4\pi\sigma = 4\pi ni$. [This value for the intensity is obtained by applying Gauss's theorem, or from analogy with a plate condenser.]

If the point considered is Q, inside the volume occupied by one of the above shells, let us suppose that that particular shell is divided into two parts of thickness $\frac{\beta}{n}$ and $\left(\frac{1-\beta}{n}\right)$ respectively, where $0 < \beta < 1$. If these are AB and CD, as in Fig. 44·14 (*b*), the current replaced by AB is βi, while that replaced by CD is $(1 - \beta)i$.

The magnetic moment of AB is $\beta i \,.\, \pi a^2$.

$$\therefore \text{ pole strength} = \pm \frac{\beta i\pi a^2}{\frac{\beta}{n}} = \pm \pi n i a^2.$$

Hence σ is as before.

Since σ has the same value for CD, we have

$$H = 4\pi ni.$$

[The last part of the above argument may be replaced by the following: let the faces of the shells be curved so that Q lies in the space between two shells. Then, by applying Gauss's theorem,

$$H = 4\pi ni$$

as before.]

An Astatic Galvanometer.—To detect a small current an astatic galvanometer may be used. Its principle of action is as follows. We have seen that a magnet tends to set itself with its axis in the magnetic meridian. If two equal magnets are arranged one above the other but with unlike poles pointing in the same direction, the resultant magnetic moment will be zero, so that the system will assume a position of rest independent of the earth's magnetic field. Since it is impossible to make and maintain two magnets of equal moment, it follows that every pair of so-called "astatic" magnets will experience a slight control due to the earth's magnetic field. [Hence it is advisable to place the instrument so that the plane of its coils lies in the magnetic meridian.] If, however, several turns of wire pass round the lower needle, as in Fig. 44·15, then when a current passes through the coils there will be greater magnetic forces on the lower magnet than on the upper one so that it will be deflected. The amount of deflexion is a measure of the current. Such instruments are practically only used to detect and not to measure small currents so that they are really astatic galvanoscopes. Frequently these effects are multiplied by arranging a coil round the upper magnet. The direction of the current in the upper coil must be different from that in the lower so that the system shall experience forces tending to urge it in one direction. A mirror, or an aluminium needle, rigidly attached to the suspended system enables the deflexions to be measured.

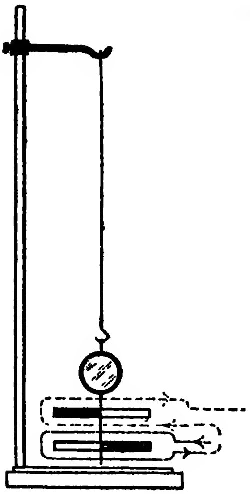


FIG. 44·15.—Astatic Galvanometer.

The Tangent Galvanometer.—Before attempting the theory of this instrument the following experiments should be performed :—A circular coil is placed with its plane at right angles to the magnetic meridian and the resultant field due to the earth's horizontal component and the current in the coil mapped with the aid of a small compass needle. A diagram similar to Fig. 44·16 (*a*) will be obtained if the magnetic field due to the current is in the direction of H_0. If, on the other hand, the coil is placed in the meridian and the resultant field mapped again, a diagram similar to Fig. 44·16 (*b*) will be obtained. This, unlike the first, is not symmetrical about the plane of the coil. If, therefore, a small magnetized needle is placed at the centre of the coil it will only be deflected in the second instance. Let us see how this deflexion enables us to measure a current.

Let m be the pole-strength of the magnet of length $2l$, while H_0 and H are respectively the horizontal component of the earth's magnetic field and the magnetic field at the centre of the circular coil due to a current i in it. If there are N turns and r is the radius of each coil, then $H = \frac{2\pi N i}{r}$. In the position of equilibrium the moments of the two couples on the magnet must be equal, i.e.

$$mH_0 \,.\, 2l \sin \theta = mH \,.\, 2l \cos \theta.$$

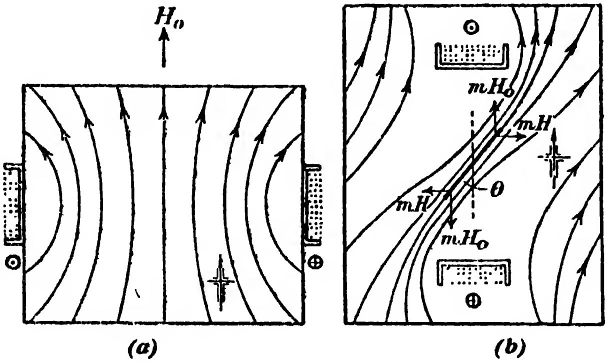


FIG. 44·16.—Principle of a Tangent Galvanometer.

Hence $H = H_0 \tan \theta$. Substituting in this expression the value of H, we have

$$i = \frac{rH_0}{2\pi N} \tan \theta.$$

In practical units this becomes

$$I = \frac{10rH_0}{2\pi N} \tan \theta.$$

This equation may be written $I = k \tan \theta$, where k is called the ***reduction factor*** of the instrument for the particular number of turns employed. This factor is not a constant since it contains H_0, which varies from place to place.

The reciprocal of $\frac{10r}{2\pi N}$, is denoted by G—it is termed the ***galvanometer constant.*** Hence

$$I = \frac{H_0}{G} \tan \theta.$$

The units for k and for G must always be stated in order to show whether or not the current is being measured in absolute or in practical units. [N.B. G is the field strength per unit current at the centre of the galvanometer coil.]

The Sine Galvanometer.—This differs from the tangent galvanometer just described in that the coil itself can be rotated about a vertical axis through its centre. In using the instrument the coil is placed in the magnetic meridian and when a current flows through the coil it is rotated until the plane of the coil contains the needle. When this occurs the field of force due to the current is at right angles to the needle. The couples on the needle are indicated in Fig. 44·17. For equilibrium we have

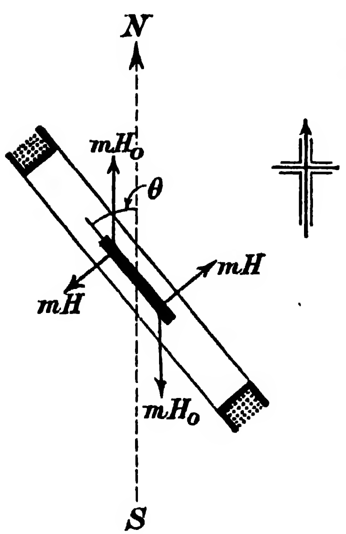


FIG. 44·17.—Principle of a Sine Galvanometer.

$$mH_0 . 2l \sin\theta = mH . 2l$$

i.e.
$$H = H_0 \sin\theta.$$

Inserting in this equation the value of H_0, we have $i = \dfrac{rH_0}{2\pi N} . \sin\theta$. Since $\sin\theta$ cannot exceed unity, the maximum current which may be measured with a sine galvanometer is $\dfrac{rH_0}{2\pi N}$.

The Sensitivity of a Tangent Galvanometer. — To determine the position on the scale where the readings will be least liable to error, let $\delta\theta$ be a small change in θ corresponding to an increment δI in the current. The error of reading the instrument will be a minimum when $\dfrac{\delta I}{I}$, the relative change in the current, is also a minimum. Since $I = k \tan\theta$, $\dfrac{dI}{d\theta} = k \sec^2\theta$, so that

$$\frac{\delta I}{I} = \frac{\sec^2\theta}{\tan\theta} . \delta\theta = \frac{2}{\sin 2\theta} . \delta\theta.$$

This expression is a minimum when $\sin 2\theta$ is a maximum, i.e. $\theta = 45°$.

The Helmholtz System of Galvanometer Coils.—It has been proved above that the magnetic intensity at a point on the axis of a circular coil carrying a current is given by the relation

$$H = \frac{2\pi i r^2}{(r^2 + x^2)^{\frac{3}{2}}}.$$

If the coil has N turns of wire, the above expression becomes

$$H = \frac{2\pi N i r^2}{(r^2 + x^2)^{\frac{3}{2}}} = \frac{a}{(r^2 + x^2)^{\frac{3}{2}}} \quad \text{[say]}.$$

The rate at which this field changes with x is given by $\dfrac{dH}{dx}$, which, for convenience, may be called y. It is important to find whether there is any region over which the above rate of change is constant, for, if this is so, by superimposing two fields due to currents in equal circular coils it should be possible to obtain a uniform magnetic field

over a considerable area. If the rate of change of $\frac{dH}{dx}$ is constant, then $\frac{dy}{dx}$ will be zero. Differentiating the expression for H with respect to x, we obtain $\frac{dH}{dx} = a\frac{d}{dx}(r^2 + x^2)^{-\frac{3}{2}} = -3ax(r^2 + x^2)^{-\frac{5}{2}}$

Differentiating again,

$$\frac{d^2H}{dx^2} = -3a[(r^2 + x^2)^{-\frac{5}{2}} - 5x^2(r^2 + x^2)^{-\frac{7}{2}}]$$

This is zero if $\frac{dH}{dx}$ is constant. Equating the above expression to zero, we have

$$5x^2(r^2 + x^2)^{-1} = 1,$$
$$\therefore x = \tfrac{1}{2}r.$$

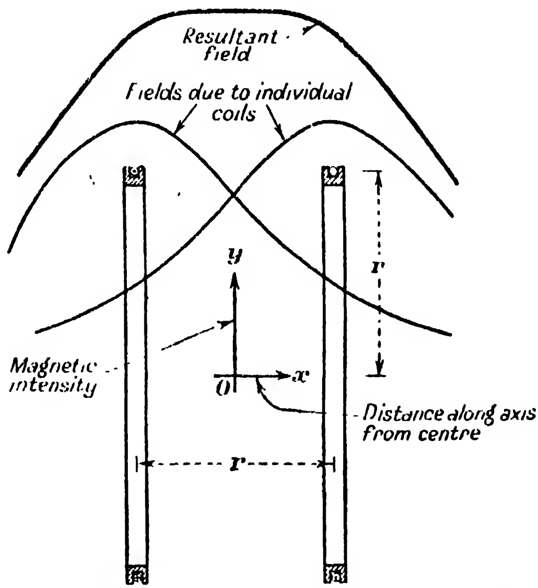


FIG. 44·18.—Helmholtz's System of Coils: Variation of Magnetic Field Strength along its Axis.

The above analysis shows that the rate of change of the field is constant at the above position. This fact is utilized in the construction of a Helmholtz tangent galvanometer. Such an instrument is not very important to-day, but the system of coils finds an important application in accurate determinations of H_0 and H_V, and also in the absolute determination of the ohm. The system consists of two coaxial coils, each of radius r and comprising the same number of turns on each. The distance between the centres of the coils is r. When the current through the coils is the same and adjusted so that the north pole of one coil faces the south pole of the other, there will be a region midway between the coils where the magnetic intensity is uniform,

for as we move along the axis from the mid-point any diminution in the intensity due to one coil is exactly compensated by the increase in the field due to the other coil. If $x = \frac{r}{2}$, we have

$$H = \frac{2\pi N i r^2}{\left[r^2+\frac{r^2}{4}\right]^{\frac{3}{2}}} = \frac{16}{5\sqrt{5}} \cdot \frac{\pi N i}{r},$$

as the numerical value of the field due to the current in each coil at the point considered. Since the coils are arranged so that the actual directions of the fields are the same, the total field at the centre of the system is double the above, i.e.

$$H = \frac{32}{5\sqrt{5}} \cdot \frac{\pi N i}{r}.$$

If the planes of the coils are in the magnetic meridian, and H_0 is the horizontal component of the earth's magnetic field, θ the angle of deflexion, $H = H_0 \tan\theta$, i.e.

$$i = \frac{5\sqrt{5}}{32\pi N} r H_0 \tan\theta.$$

The manner in which the field due to each coil varies and the region over which the combined field is uniform is shown in Fig. 44·18.

Experiment.—The manner in which the magnetic field due to a circular current varies at points along its axis may be investigated experimentally by placing a coil with its axis normal to the magnetic meridian. The magnetic field strength at a point on its axis is proportional to the tangent of the angular deflexion of a magnetometer needle placed with its centre at that point. A graph exhibiting the relation may then be drawn.

The Schuster Magnetometer.—The most accurate method of determining the horizontal component of the earth's magnetic field was proposed by SCHUSTER in 1914. The actual research was carried out by F. E. SMITH and completed in 1923. The principle of this instrument, which has been termed the Schuster magnetometer, is as follows :—Two equal coils are arranged at a distance apart equal to their common radius as in the Helmholtz galvanometer. A small magnet is suspended by a quartz fibre so that its centre is on the axis of the coils and midway between them. When a current is sent through the coils in the same direction a magnetic field is produced which is uniform over a considerable region in the neighbourhood of the magnet. Suppose that the planes of the coils are normal to the magnetic meridian. Then the magnetic field due to the current in them is parallel to the direction of H_0. If the sense of the field is the same as that of H_0, then the magnet remains undeflected for all values of the field due to the current. If the current through the coils is reversed, the sense of the magnetic field due to it will be opposite to that of H_0. Let G be the field per unit current (e.m.u.) at the centre of the coil system due to a current in the coils. If the current is i, the magnetic field is Gi. As long as Gi is less than H_0, the magnet will continue to point to the magnetic north, but when Gi exceeds H_0 the magnet swings round through 180° so that its north pole now points to the south. It continues to do this for all values of $Gi > H_0$. If, however, the coil system is rotated through

a small angle—its magnitude depends on the difference between Gi and H_0, so that this should be as small as possible—then the magnet may be made to set at right angles to the meridian. Suppose that these conditions have been realized.

Let NS, Fig. 44·19, be the suspended magnet so that when deflected it is at right angles to H_0. Let the normal to the plane of the coils make an angle α with OK. Let m be the pole strength of the magnet and i the current in the coils. Then the magnet is in equilibrium under the action of the two couples indicated. Since the moments of these are equal when equilibrium has been reached, we have

$$mH_0 . 2l = mGi . 2l \cos \alpha$$

if $2l$ is the length of the magnet. Hence

$$H_0 = Gi \cos \alpha.$$

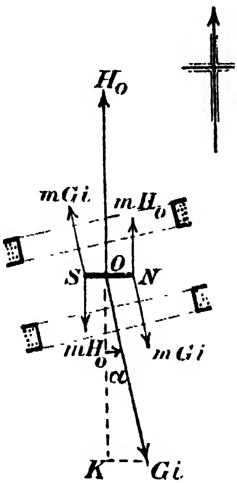


FIG. 44·19.—Schuster Magnetometer.

The value of the current was adjusted so that α was small when NS was in the desired position. The current was determined by 'weighing' it with the aid of an ampere balance. The coils themselves consisted of twelve turns of bare copper wire wound on a marble cylinder of radius 30 cm. The suspended magnet was about 1 cm. long and 5 mm.² in cross-section. This was supported by a quartz fibre 25 cm. long carrying a reflecting mirror and damping vane.

Finally H_0 was determined with an error of 3 parts in 10^5, the actual observations taking 4 minutes. The method is exceptionally good since it is rapid and sensitive, and errors such as non-uniformity of magnetic field, possible magnetic effects of the material of the coil supports, and possible electrostatic effects on the suspended system were eliminated by paying special attention to the design of the magnetometer.

Bates' Apparatus for the Measurement of the Horizontal Component of the Earth's Magnetic Field.—This is an adaptation of the Schuster-Smith magnetometer for student's use. In the original method two coils were arranged as in a Helmholtz galvanometer to produce a uniform magnetic field to deflect a suspended magnet through an angle $\frac{\pi}{2}$, when the above field made a small angle, α, with H_0. In the present apparatus only one coil was used.

In order to measure a rotation of $\frac{\pi}{2}$ accurately, BATES used the apparatus depicted in Fig. 44·20 (a). It was very satisfactory. The small magnet ns was suspended by a thin fibre of unspun silk at the centre of a vertical coil of wire AB. The normal to the plane of this coil must make a small angle α with the axis of ns in its undeflected position. A small plane mirror, M, was attached below this magnet, so that the plane of the mirror made an angle of about 45° with the axis of the magnet. Two plumb lines were placed at X and Y, so that

they were in a vertical plane containing the fibre. [It is not essential for XY to be perpendicular to *ns*.] For convenience, two sheets of white cardboard, C and D, were placed behind the plumb lines as indicated.

Let us suppose that the light from an illuminated slit S, after traversing the convex lens L, was reflected from M and that the lens was so adjusted that an image of the slit was produced on C. By suitably adjusting the positions of S and L, this image was caused to produce a shadow of X which bisected the image. If now the magnet rotated through $\frac{\pi}{2}$, the beam of light reflected by M turned through π, and the image of S was now bisected by the shadow of Y produced by it. The line XY need not be normal to the magnetic meridian, in fact, it must be chosen so that AB does not interfere with the light.

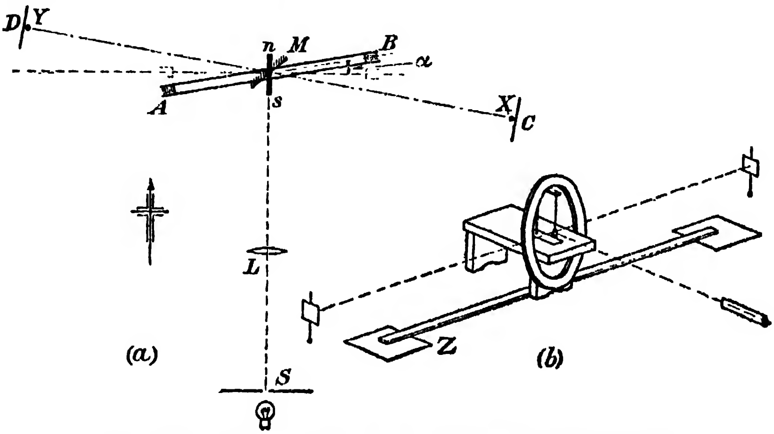


FIG. 44·20.—Bates' Apparatus for the Measurement of H_0.

A simple form of apparatus which has been found suitable for rapid work is shown in Fig. 44·20 (*b*). The magnet system is suspended from a torsion head fixed in a brass holder, so that the magnet lies at the centre of the coil. The torsion head is not necessary if a very long thin fibre and a strong magnet are used. A stop is provided so that the system may not turn through an angle greater than π. The arrangement of the lamp and plumb lines is clearly shown in the diagram. A box with windows, for shielding the system from draughts, is not shown.

In order to prepare the apparatus for use, the magnet must be removed from the brass tube holding it, and the torsion head turned until the tube above sets approximately along the magnetic meridian, and perpendicular to the plane of the coil. The magnet is then replaced, and a small current passed through the coil. In general, the magnet will be deflected, and the whole apparatus must be rotated until a position is found where no deflexion is produced. The plane of the coil is then at right angles to the magnetic meridian, and the small current merely serves to assist or reduce the effect of the horizontal component of the earth's magnetic field upon the magnet.

The whole apparatus is then rotated, say, anti-clockwise, so that the coil moves through a small angle α about a vertical diameter. To measure this angle, a metre scale Z is suitably attached to the base of the apparatus. The ends of this rod lie immediately above graph paper. The distances through which the ends of the rod move when the apparatus is rotated are recorded. If these are d_1 and d_2 respectively; then α is equal to $(d_1 + d_2)/100$ radians.

The lamp, etc., are then adjusted so that the light reflected from the mirror falls upon X. A current, measured by a potentiometer method, is passed through the coil and gradually increased until the light falls on Y. If the torsion in the fibre is large the torsion head should be rotated through $\frac{\pi}{2}$ in such a direction that the twist in the suspension is reduced to zero. More accurately, the rotation should be $\left(\frac{\pi}{2} + \alpha\right)$. We then have

$$H_0 = \frac{2\pi NI}{10r} \cos \alpha = \frac{2\pi N i_2}{r} \cos \alpha,$$

where I is the current in amperes, N the number of turns of wire in the coil, and r the effective radius of the coil.

$\frac{2\pi N}{r}$ is the G of the previous section, and we see that G has dimensions $cm.^{-1}$

ELECTRODYNAMICS

The Mechanical Force on Currents in a Magnetic Field.—Since when a magnet is introduced into a magnetic field it experiences mechanical forces, the fact that a current in a closed circuit is equivalent to a magnetic shell naturally causes us to expect that when a conductor carrying a current is introduced into a magnetic field it will experience a mechanical force.

The following experiment, due to FARADAY (1822), shows the existence of this mechanical force. A, Fig. 44·21, is a glass tube provided with a close fitting cork C through which passes a cobalt steel magnet NS. The cork is covered with mercury. X is a piece of copper wire free to rotate about a pivot P connected to one pole of a battery B. The lower end of the wire X dips into the mercury. The other electrode of B is connected to the mercury by a wire passing through C. Under these conditions a current passes down the wire X as indicated and this will be found to rotate in a clockwise direction as seen from above. If the current is reversed, the wire will move with the same angular velocity in the opposite direction.

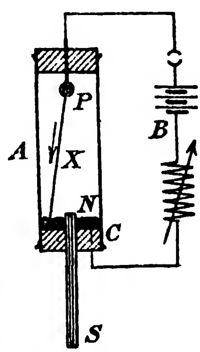


FIG. 44·21. — Faraday's Experiment to show the Force on a Conductor in a Magnetic Field.

It has been shown that the magnetic field strength at a point P at a distance r from a straight wire carrying a current i is $2i/r$. If a magnetic pole of strength m is placed at P it will experience a force $2im/r$ in the direction indicated in Fig. 44·22. Since action and reaction are equal and opposite it follows that in consequence of this force on m, there will be an equal and opposite force on the wire. Its direction is shown in the diagram.

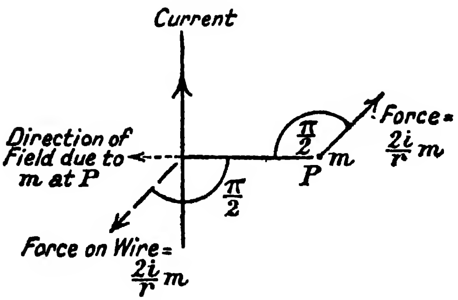


FIG. 44·22.

From the above it will be seen that the direction of the force on a conductor carrying a current in a magnetic field is expressed by ***Fleming's Left-Hand Rule.*** According to this, if the thumb and first two fingers of the left hand are extended so that they are at right angles to one another, and the first finger points in the direction of the magnetic field, the second in the direction of the current, then the thumb points in the direction of the mechanical force on the conductor. This is illustrated in Fig. 44·23.

[Strictly speaking, this rule only applies to the particular instance when the wire is in air and $\widehat{iH}$, i.e. the angle between the directions of i and H is $\frac{\pi}{2}$. In other instances the first finger must point in the direction of that component of H which is at right angles to the wire and in the plane containing i and H.]

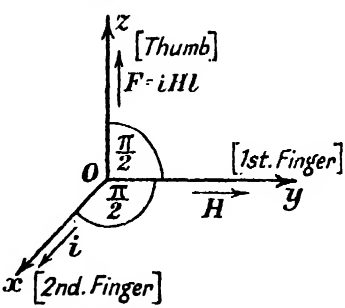


FIG. 44·23.—Fleming's Left-Hand Rule.

If l is the length of a straight conductor carrying a current i [e.m.u.] and H is the intensity of the magnetic field, the magnitude of the force on the wire is given by $F = iHl$.

Barlow's Wheel.—This experiment is another illustration of how electrical energy may be converted into mechanical energy as in an electric motor and also enables one to verify Fleming's left-hand rule. A copper wheel, Fig. 44·24, is supported on a horizontal brass axle as indicated. The supports for the axle are carried on ebonite pillars. The periphery of the wheel makes

contact with a pool of mercury, A, placed in the wooden base of the instrument. A smaller copper wheel fixed to the axle just dips into another pool of mercury, B. These mercury pools are connected to a battery so that a current passes from the axle to the periphery of the wheel. A powerful horse-shoe magnet is placed so that the lines of force are perpendicular to the plane of the wheel over a considerable portion of it. When the current is passing the wheel rotates in a direction given by Fleming's Left-hand Rule, viz. when the current passes from the axle to the periphery of the wheel, the motion at A is away from an observer to whom the lines of magnetic force run from right to left.

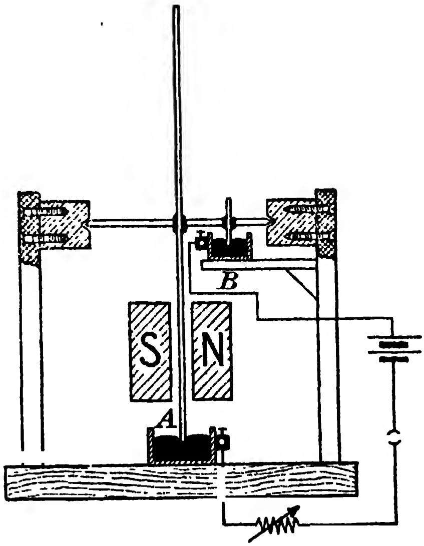


FIG. 44·24.—Barlow's Wheel.

The Force on a Conductor carrying a Current in a Magnetic Field.—Let P, Fig. 44·25, be a point at distance a from a long straight conductor carrying a current i. Then the magnetic intensity at P is $\frac{2i}{a}$ [cf. p. 816]. This means that if a unit positive pole is placed at P, it will experience a force $\frac{2i}{a}$—its direction is indicated. There will be an equal and opposite force acting on the wire due to the unit magnetic pole at P. This will be equal to the resultant of all the forces acting on the different elements of length into which the conductor may be supposed divided. Let AB be such an element of length δs, where s is the distance of A from O, the projection of P on the wire. Then H, the magnetic field strength at A (the medium is assumed to be air), due to the unit pole at P is $\frac{1}{r^2}$, and is directed along PC, where C is the mid-point of AB. Let the force, δF, on this element be

$$H \, . \, f(\theta) \, . \, i \, . \, \delta s,$$

where $f(\theta)$ is to be determined. It

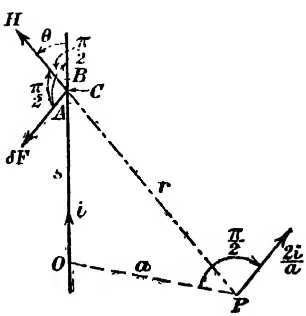


FIG. 44·25.— Force on a Conductor.

will be assumed that this force is parallel to the resultant force on the wire. Then the total force on the wire due to unit magnetic pole at P is

$$2\int_{\frac{\pi}{2}}^{0} \frac{1}{r^2} . f(\theta) . i . ds = \text{F [say]}.$$

But $r^2 = a^2 + s^2$, and $s = a \cot \theta$, i.e. $ds = -a \operatorname{cosec}^2 \theta . d\theta$.

$$\therefore \text{ F} = -2\int_{\frac{\pi}{2}}^{0} \frac{1}{a^2 \operatorname{cosec}^2 \theta} . f(\theta) . i . a \operatorname{cosec}^2 \theta . d\theta$$

$$= -\frac{2i}{a}\int_{\frac{\pi}{2}}^{0} f(\theta) . d\theta$$

But F, the resultant force on wire, $= \frac{2i}{a}$

$$\therefore \int_{\frac{\pi}{2}}^{0} f(\theta) d\theta = -1.$$

This is satisfied if $f(\theta) = \sin \theta$.

$\therefore$ Force on element due to the unit magnetic pole at P $= \frac{i \sin \theta \delta s}{r^2}.$

The force on AB is therefore $i\text{H} \sin \theta . \delta s$, since H, the field at A due to the unit pole at P, is $\frac{1}{r^2}$.

If the magnetic field is uniform and everywhere normal to a wire of length l carrying a current i, the force on the wire is $li\text{H}$.

Laplace's Law.—It has just been established that the force on an element δs of a straight conductor carrying a current i is

$$\frac{i \sin \theta . \delta s}{r^2},$$

where r is the distance of the element from a unit magnetic positive pole. Since action and reaction are equal and opposite, it follows that the magnetic field strength at P due to the current in the wire is

$$\frac{i \sin \theta . \delta s}{r^2}.$$

This is LAPLACE's law.

If there is a pole of strength δm at P, the force on it due to the current in δs is

$$\frac{\delta m . i \sin \theta . \delta s}{r^2}.$$

The Mutual Action of Currents.—AMPÈRE first investigated the action between wires carrying currents. We shall limit our discussion to parallel wires. He found that when the currents flowed in the same direction there was an attraction between them. On the other hand, when the currents flowed in opposite

directions there was repulsion. The lines of force in a horizontal plane when the currents are passing vertically downwards are given in Fig. 44·26 (*a*). We notice that all the lines (tubes) of force surround both conductors, and since there is a tension along a tube of force these will tend to contract and draw the wires together. The corresponding field when the currents pass in contrary directions is given in Fig. 44·26 (*b*). In this instance no tube of force surrounds ***both*** conductors, and since the tubes are more crowded together in the region between the wires, these will be pushed aside in virtue of the lateral thrust which exists along a tube of force.

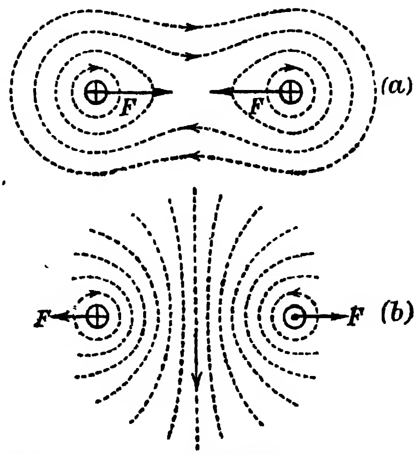


FIG. 44·26.—Lines of Magnetic Force due to Currents in Parallel Wires.

Experimental Illustrations.—In Fig. 44·27 there is represented a long coil of copper wire about 6 cm. in diameter. Its upper end rests in mercury while its lower end just touches some mercury in

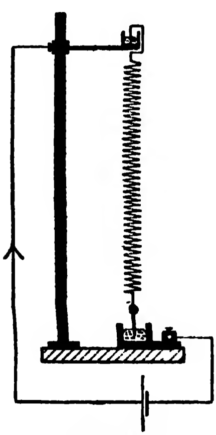


FIG. 44·27.

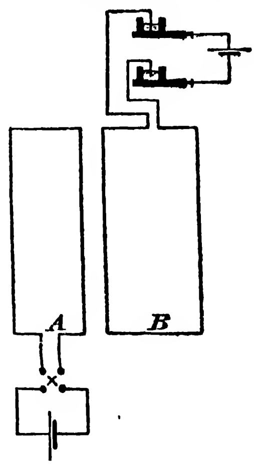


FIG. 44·28.

a second container insulated from the first. When a current is passed through the coil the mutual attractions between its various turns causes the coil to contract: the current is broken and the

coil expands, thereby completing the circuit again. The process is then repeated.

Fig. 44·28 indicates a fixed coil A and a movable one B. The two coils are not in the same plane. When a current passes through each coil attraction ensues if the currents in the adjacent sides pass in the same direction. There is repulsion when one of the currents is reversed.

Fig. 44·29 (*a*) depicts a coil of wire [a solenoid] attached to a piece of wood so that it shall be rigid. Its two ends dip into mercury

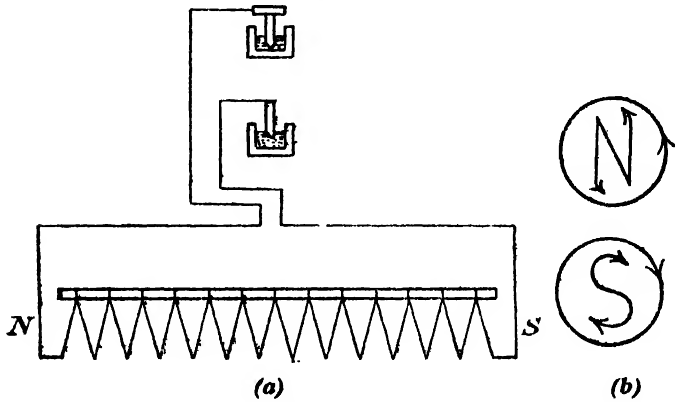


FIG. 44·29.

cups so that the coil is free to rotate about a vertical axis. By making contact between the cups and a battery a current may be passed through the solenoid. Let us assume that when an observer looks along the coil in the direction SN, the current appears to flow in a clockwise direction. The magnetic field at that end will possess south-seeking polarity. An aid for memorizing this is shown in Fig. 44·29 (*b*). The end S of the coil will therefore be attracted by the north-seeking pole of another magnet. It will also be attracted by another solenoid if the current in the latter flows in the appropriate direction.

Suspended Coil.—The problem of determining the forces on a rectangular coil suspended in a magnetic field is very important since upon its solution depends the principle of many current measuring instruments. Let ABCD, Fig. 44·30, be a fixed rectangular coil of length l and breadth b placed in a magnetic field of strength H, the direction of the field always being in the plane of the coil. [This field is a radial one—it is obtained by using the system shown in Fig. 44·32 (*b*).] If the current flows in the direction indicated each vertical side of the wire will experience a force iHl

acting in the directions shown. These constitute a couple whose moment is $iHl \cdot b = iaH$, where a is the area of the coil.

If the coil is made of N turns of wire each with an area a, the moment becomes $N \cdot iaH = iAH$, where $A = Na$, the effective area of the coil.

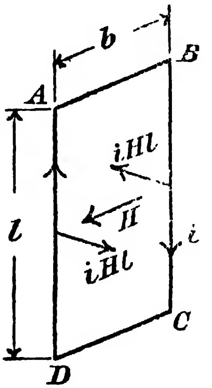


FIG. 44·30.—Couple on a Rectangular Coil carrying a Current in a Radial Magnetic Field.

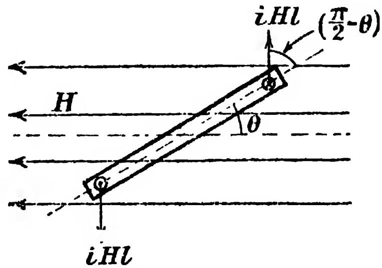


FIG. 44·31.—Couple on a Rectangular Coil carrying a Current in a Uniform Magnetic Field.

Now suppose that the coil is suspended in a uniform magnetic field parallel to the zero position of the coil—cf. Fig. 44·31. If the coil is deflected through an angle θ, each force iHl is inclined to the plane of the coil at an angle $\left(\frac{\pi}{2} - \theta\right)$. The couple is therefore $iHl \cdot b \cos \theta$. If there are N turns of wire, the couple is $iHA \cos \theta$, where $A = Nlb$. For equilibrium

$$iAH \cos \theta = c\theta,$$

where c is the restoring couple due to the suspension when the twist in it is one radian.

The Suspended Coil Galvanometer.—This is a most reliable and sensitive instrument for detecting electric currents. A narrow coil, Fig. 44·32 (a), consisting of many turns of wire (only the frame on which these are wound is shown) is suspended from a movable head by a fine phosphor bronze wire between the poles of a strong magnet. When a current flows through the coil, forces act in contrary directions on the opposite sides of the coil, i.e. a couple acts on the coil which rotates until the twist in the suspension produces an equal and opposite couple. To make the couple acting on the coil as large as possible for a given current the coil must consist of many turns of wire and be placed in a strong magnetic field. To increase the deflexion the suspension is made very fine so that the torsional couple per unit angular displacement is small. In order to concentrate the field and arrange that

it is always parallel to the plane of the coil for all positions of the latter not far removed from its position of rest a cylindrical piece of soft iron is fixed midway between the poles of the magnet. The current enters the coil through the suspension and leaves through a very fine spiral below the coil. This spiral is made of very fine wire and consists of several relatively large turns so that it shall exert only a small restoring couple on the coil. A mirror

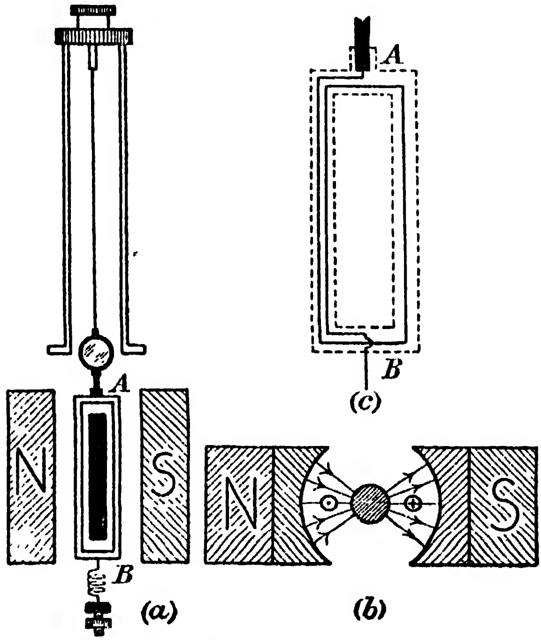


FIG. 44·32.—Suspended Coil Galvanometer.

rigidly attached to the framework of the coil or to the lower and thicker portion of the suspending wire—there must be no relative motion between the mirror and the coil—enables any small deflexion to be measured. Let us assume that θ is the angle of deflexion and that c is the couple in the wire due to unit (radian) difference of twist between its ends. Then $c\theta$ is the couple in the present instance. But we have seen that the moment of the forces on the coil is iAH. When equilibrium has been attained this must equal the couple due to the twist in the suspension, i.e.

$$i\mathrm{AH} = c\theta, \text{ or } i = \frac{c\theta}{\mathrm{AH}}.$$

Dead Beat and Ballistic Galvanometers.—The theory given above applies only to steady currents. In moving-coil galvano-

meters designed to measure only steady currents the coil is wound on a metal frame (copper): this is highly damped, i.e. it is very quickly brought to rest, when it moves in a strong magnetic field [cf. p. 950] and the instrument is ***dead-beat***. If, however, the coil is wound on a celluloid or a cane frame, the instrument is not dead-beat unless the external resistance is less than a certain critical value—it is ***ballistic***, i.e. it does not attain its final position at once but oscillates about it, the amplitude of the oscillations gradually decreasing. Although it may still be used to measure a steady current, its real value lies in the fact that it can detect transient currents, i.e. currents which exist for a short but finite time. It actually measures the quantity of electricity which passes through the coils, e.g. the discharge from a condenser. In order to do this an essential feature of a ballistic galvanometer is that the moving system shall not have moved from its zero position before the whole of the quantity to be measured has passed. To ensure this the moment of inertia of the system about its axis of suspension must be as large as possible, i.e. the time of swing must be large. Also, the couple tending to restore the system to its equilibrium position when it is displaced must be as small as possible.

Thus if a condenser is charged and then discharged through a ballistic galvanometer, the system will move in the above manner owing to the impulse it has received due to the passage of a quantity of electricity through it. It may be shown that the magnitude of the first swing outwards [the deflexion of the galvanometer, its 'throw' or 'kick'] is proportional to this quantity of electricity (Q).

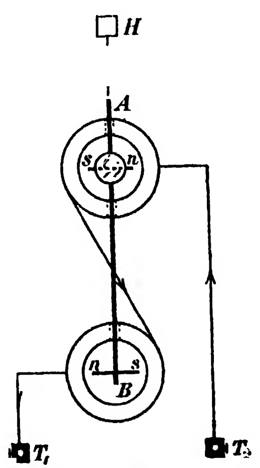


Fig. 44·33.—Moving Magnet Ballistic Galvanometer.

Kelvin's Moving Magnet Galvanometer.—In principle a moving magnet galvanometer is a tangent galvanometer, only it is much more sensitive. The formula for a tangent galvanometer is

$$i = \frac{r\mathrm{H}_0}{2\pi\mathrm{N}} \tan \theta.$$

From this it appears that in order to measure a small current, r must be small, and N large, for under these conditions θ becomes larger. It is impossible, however, to make r small and at the same time have a large number of turns unless the diameter of the wire is very small. This

increases the resistance of the instrument, a condition not always desirable. Hence, in practice, a compromise must be effected. A type of moving magnet galvanometer designed by Kelvin is shown in Fig. 44·33. It uses fixed coils and a moving system of magnets, the axis of the coils being normal to that of the magnet system. The magnet system is an astatic one: this is used so that the restoring couple on the system shall be small—an essential condition if great sensitivity is to be obtained. The highest sensitivity is obtained by adjusting a controlling magnet so that the magnet system lies in a field which is small but not quite uniform. This last condition is only necessary if the system is perfectly astatic—very seldom if ever obtained in practice. AB is an aluminium or glass rod suspended by a quartz or silk fibre. The magnets are such that their planes are accurately parallel, but their polarities are reversed.

The coils carrying the current are wound in contrary directions so that the couples on the upper and lower magnets assist each other.

These galvanometers are considerably affected by stray non-uniform magnetic fields, so that they are generally screened by an iron shield cylindrical in shape and surrounding the instrument.

Kelvin galvanometers may be used to measure either transient or steady currents, the deflexion being noted with the aid of a mirror rigidly attached to the magnet system.

Very often each magnet is replaced by a system of three short magnets; in this way the moment of inertia of the system about its axis of rotation is reduced while the effective pole strengths are increased. Both these conditions are desirable.

Comparison of the Moving Magnet and Suspended Coil Galvanometers.—

Moving Magnet Type	Suspended Coil Type
(i) May be used to measure transient currents as well as steady currents.	May be used to measure transient currents as well as steady currents.
(ii) The system is always ballistic.	The system is ballistic if the frame on which the coil is wound is made of ivory, cane, etc.
(iii) H is varied by means of a control magnet fixed to the instrument. The field due to this magnet is generally arranged so that the horizontal component of the earth's magnetic field is diminished: the instrument is then more sensitive.	H is fixed.

Moving Magnet Type.	Suspended Coil Type
(iv) The motion is not affected by short circuiting the coils.	The system may be brought very quickly to rest by short circuiting the coils, for this then forms part of a closed circuit moving in a very strong magnetic field, and the motion is retarded by the effects of the induced e.m.f.
(v) When used to measure the charge on a condenser the throw is independent of the resistance in the circuit—unless the resistance exceeds 10^6 ohms when the rate at which the discharge takes place is slowed down and the magnet moves before the discharge is completed.	When used to measure the charge on a condenser the throw is independent of the resistance in the circuit—unless the resistance exceeds 10^6 ohms when the rate at which the discharge takes place is slowed down and the magnet moves before the discharge is completed.
(vi) When measuring the charge induced in a coil it must be remembered that the charge is inversely proportional to the resistance of the circuit, so that the galvanometer throw also varies in the same way. This is only true providing that the resistance is not so large that the time for the discharge to take place becomes appreciable in comparison with the period of the galvanometer.	The suspended coil galvanometer cannot be used if the resistance of the circuit is so low that the motion of the coil is not ballistic. In such instances it is necessary to place a resistance in series with the galvanometer, and of such a value that the motion is ballistic. This reduces the quantity of electricity passing for an induced e.m.f. of given magnitude. Moreover, the resistance must not be so large that the time of discharge becomes comparable with the period of the instrument.
(vii) The instrument may be used in any position with reference to the earth's magnetic field, a control magnet effecting any desired orientation of the needle.	The instrument is not affected by external magnetic fields of ordinary magnitudes.
(viii) The time of swing may be changed by altering the position of the control magnet.	The time of swing is fixed.
(ix) The instrument must be screened from external variable magnetic fields.	No screening is necessary.
(x) The needle may be brought to rest by placing a solenoid near the galvanometer. This is connected to a cell and tapping key. The key is momentarily depressed when the swing is in such a direction that by so doing the amplitude is reduced considerably.	The coil may be brought to rest by short circuiting it.

EXAMPLES XLIV

1.—Define the absolute and practical units of potential difference and current. How is the electrical resistance of an electric circuit defined? In what units is it measured?

2.—Indicate how the magnetic intensity due to a linear current may be calculated from a consideration of the equivalent magnetic shell.

3.—Describe (*a*) the Daniell cell, (*b*) the Leclanché cell, and give an account of the processes which take place in each when it is in action.

4.—Define the electromagnetic unit of current and state what relation it bears to the ampere. A circular coil of 10 turns and 10 cm. diameter is placed in the magnetic meridian and has a small magnet at its centre. Calculate the current in amperes which will deflect the needle 45° if the horizontal component of the earth's field is 0·2 oersted.

5.—What is a uniform magnetic shell? Define the strength of such a shell.

Derive an expression for the magnetic intensity at a point in the middle of a solenoid 40 cm. long and 1 cm. in radius, wound with 400 turns of wire carrying a current of 5 amperes.

6.—Derive an expression for the magnetic potential at a point in air due to a short bar magnet, and deduce an expression for the magnetic potential due to a uniform magnetic shell. Apply this result to calculate the magnetic intensity at a point due to a linear current of 2 amp. in its neighbourhood.

7.—What precautions must be taken to obtain an accurate value for the deflexion of the needle of a tangent galvanometer when a current is passed through its coils?

If H_0 is the controlling field in which a tangent galvanometer is placed, show that the resultant horizontal field when the needle is deflected through an angle ϕ is H_0 sec. ϕ.

It is observed that when a certain current is passed through a given tangent galvanometer that the needle is deflected 60°, and that when disturbed it oscillates about its position of equilibrium at a rate of 20 cycle.min.$^{-1}$ When the current is reduced to give a deflexion of 45°, what is the rate of oscillation of the needle about its new position of equilibrium?

CHAPTER XLV

OHM'S LAW AND ITS APPLICATIONS

The E.M. Unit of Current.—We have already defined this unit of current and, basing our argument on this definition, shown that the magnetic intensity at the centre of a coil of radius r and carrying a current i (e.m.u.) is $\frac{2\pi i}{r}$. Students who find the previous argument difficult may therefore assume this result and define the electromagnetic unit of current as follows:—it is *that steady current which, when flowing in a circle of unit radius, produces at its centre a magnetic field of strength 2π oersted,* or *if the circle has a radius r, the field at the centre is $\frac{2\pi}{r}$ oersted.*

The Practical Unit of Current.—For many purposes, the above unit is too large, so that the practical unit of current is defined as one-tenth of the e.m.u. of current. It is called the *ampere.* Thus, 10 amperes = 1 e.m.u. of current.

The Electromagnetic Unit of Quantity of Electricity.—This is defined as *the amount of electricity flowing per second through a conductor which is carrying a steady current of one e.m.u.* It is sometimes called a *weber.*

The Practical Unit of Quantity of Electricity.—This is termed the *coulomb* or *ampere-second,* and is the amount of electricity flowing per second through a conductor when the current in it is one ampere. Thus 10 coulombs = 1 weber.

The International Ampere.—The ampere already defined is the *true ampere:* to provide a convenient working definition of the practical unit of current the chemical effects of a current are utilized. *The international ampere is that unvarying current which when passed through an aqueous solution of silver nitrate deposits silver at the rate of 0·001118 gm. sec.*$^{-1}$ Consequently, the *international coulomb* is that amount of electricity which will deposit 0·001118 gm. of silver from an aqueous silver nitrate solution.

The international ampere was intended to be equal to the true ampere: actually it is 0·025 per cent. smaller. In ordinary practice this slight difference is neglected.

Electromotive Force and Potential Difference.—We have seen that when a copper and a zinc rod are dipped into dilute sulphuric acid a current flows from the copper to the zinc when these are connected by a wire. This is because as soon as the plates are placed in the acid there is established between them a potential difference [p.d.] with respect to the liquid. The copper is positive and the zinc negative. The reason for this is that while both metals have a tendency to pass into solution and carry positive electricity with them, the tendency is greater with zinc. Hence an excess of zinc atoms with positive charges—termed ions [*ἰόν* a wanderer] —pass into the solution and leave the zinc negatively charged. The copper, on the other hand, acquires a positive charge. The motion of the positive electricity through the cell is due to a chemical electromotive force (e.m.f.). Thus, the e.m.f. of a cell acts from the zinc to the copper and drives positive electricity to the copper. It is in virtue of this e.m.f. that there is established between the two metal plates a difference of potential. This potential difference does not increase indefinitely since there is only a finite e.m.f. in the cell. It only rises until the tendency for positive electricity in the cell to move towards the copper under the influence of the e.m.f. is neutralized by its tendency to move towards the zinc under the influence of the p.d. between the plates. Thus, when no current is passing through the cell its e.m.f. is equal to the p.d. between its plates.

An e.m.f. and a p.d. are measured in the same units.

The e.m.u. of Potential Difference.—*Suppose that A and B are two points in a conductor through which a current is flowing. Let this current flow for such a time that one e.m.u. of quantity of electricity passes across each section of the wire normal to the lines of flow of the current. Then if the energy liberated is one erg, the potential difference between A and B is one e.m.u. of potential.*

The Practical Unit of Potential Difference.—This is termed the *volt. If the current flows for such a time that one coulomb passes across each section of the conductor normal to the lines of flow of the current and the energy liberated is one joule [the practical unit of energy], the potential difference between A and B is one volt.*

Relation between the Electromagnetic and the Practical Units of Potential Difference.—If a definite p.d. exists between

two points in a conductor, the same amount of energy will always be liberated irrespective of the units used to express the quantity of electricity and the potential difference. Now when one coulomb is transported across each section of a conductor between whose ends there is a potential difference of one volt, the energy liberated is one joule, or 10^7 ergs. Since one e.m.u. of quantity of electricity $\equiv$ 10 coulombs, the energy liberated when this quantity of electricity passes each cross-section of the above conductor is $10 \times 10^7 = 10^8$ ergs. But since, when one weber is transported across each section of a conductor between whose ends there is a potential difference of one e.m.u. of potential, the work done is one erg, it follows that

$$1 \text{ volt} \equiv 10^8 \text{ e.m.u. of potential.}$$

[Strictly speaking, this is the ***true volt.***]

Ohm's Law.—***When a steady current is flowing through a conductor the potential difference between its ends divided by the current is a constant, provided that the physical condition of the conductor does not change.*** This constant is termed the ***resistance*** of the conductor. It is measured in true ***ohms*** when the potential is in true volts and the current in true amperes.

A conductor has a resistance equal to one e.m.u. of resistance if the p.d. between its ends is one e.m.u. of potential when the current through it is one e.m.u. of current. Thus

$$1 \text{ true ohm} = \frac{1 \text{ true volt}}{1 \text{ true ampere}} = \frac{10^8 \text{ e.m.u. of potential}}{10^{-1} \text{ e.m.u. of current}} = 10^9 \text{ e.m.u.}$$
of resistance.

The reciprocal of the resistance of a conductor is termed its ***conductance.*** The practical unit of conductance is the ohm^{-1}

The International Ohm.—This is defined as ***the resistance of a column of mercury, at the temperature of melting ice, 14·4521 gm. in mass, of constant cross-sectional area, and of length 106·300 cm.***

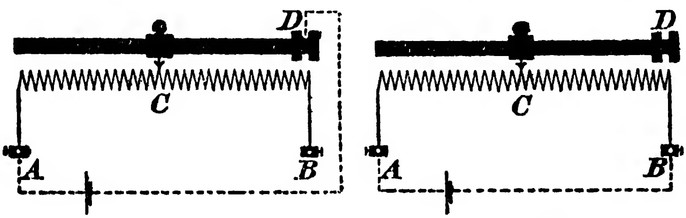
The international ohm was intended to be a practical realization of the true ohm or 10^9 e.m.u. of resistance. Actually it is slightly larger.

The Evaluation of the Ohm.—The system of 'practical' units was devised originally by a committee of the British Association: the value of the true ohm was determined experimentally, and standard resistance coils (German-silver) were constructed (1863). In 1881 the first International Congress of Electricians advocated a redetermination of the ohm in absolute measure:

wires underneath the top of the box to the points A and B respectively. One of the leads from the battery and one from the galvanometer are then connected to H and K respectively so that when the keys are depressed there is connection between these leads and A and B as in the first form.

Very frequently it is stated that the battery and galvanometer may be interchanged. Theoretically this may be done, but in practice, with the bridge arranged as in the above numerical example, there would be a relatively large current through Q and S, and a much smaller one through P and R. This may cause Q and S to be heated considerably and thereby alter their resistances. Care must therefore always be taken to see that the bridge is arranged so that only small and nearly equal currents flow through the various arms.

Adjustable Resistors.—To vary the current in a circuit use is made of a variable resistor or rheostat. This may consist of



(a) Adjustable Sliding Resistor.

(b) Adjustable Resistor used as a Potential Divider.

FIG. 45·8.

a number of carbon plates held together in a suitable frame, the resistance being reduced by applying pressure by means of a screw. Another form of variable rheostat is represented in Fig. 45·8 (*a*). AB is a wire wound on a frame and C is a movable contact carried on a rod of triangular section. If a cell is connected to A and D the current flows through the portion AC of the rheostat. By moving the sliding contact to the left the resistance in the circuit is diminished. Such a resistance may be used as a potential divider. For this purpose a battery is connected across AB so that a current flows through the whole resistance—cf. Fig. 45·8 (*b*). This establishes a potential difference between B and C, and therefore between B and D. When C is moved to the left this potential difference increases.

Kirchhoff's Laws.—These are two rules which enable us to solve problems concerning currents flowing in a network of wires. They state:

(*a*) ***In any network of wires the algebraic sum of the currents which meet at a point is zero,*** i.e. $\Sigma I = 0$.

(b) The algebraic sum of the electromotive forces in any closed circuit or mesh is equal to the algebraic sum of the products of the resistances of each portion of the circuit and the currents flowing through them, i.e. $\Sigma E = \Sigma IR$.

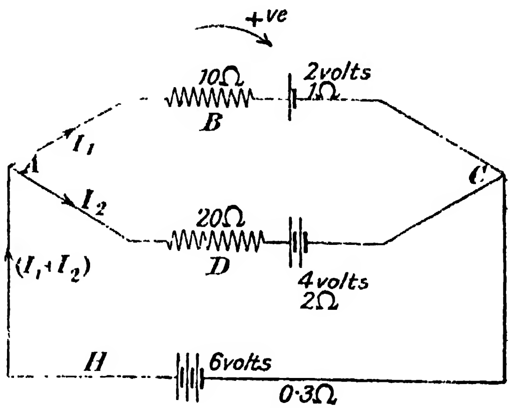


FIG. 45·9.

Example.—The arm ABC of a circuit contains a resistance of 10 ohms and a cell of internal resistance 1 ohm and e.m.f. 2 volts, while the branch ADC contains a resistance of 20 ohms and two similar cells. Across AC there is placed a battery of e.m.f. 6 volts with an internal resistance of 0·3 ohm. Calculate the currents through the two resistances if the e.m.fs. are directed as in Fig. 45·9.

Let I_1 and I_2 be the currents in ABC and ADC, so that the current in CHA is $(I_1 + I_2)$—by Kirchhoff's first law applied to the point A. Then considering the mesh ABCD and taking an e.m.f. to be positive when it acts round the mesh in a clockwise direction, we have

$$10I_1 + 1 \, . \, I_1 - 2I_2 - 20I_2 = -2 + 4$$

i.e. $$11I_1 - 22I_2 = 2.$$

Similarly from the mesh ADCH, we have

$$20I_2 + 2I_2 + (I_1 + I_2)\, 0{\cdot}3 = -4 + 6$$

i.e. $$22{\cdot}3I_2 + 0{\cdot}3I_1 = 2.$$

Solving these equations $I_1 = 0{\cdot}35$ amp. and $I_2 = 0{\cdot}085$ amp.

Maxwell's Cyclic Currents.—The above method of determining the current in any part of a circuit becomes complicated when the circuit has many branches. Maxwell suggested the 'cyclic current' device to simplify the problem. He imagined that a specified cyclic current flowed in each mesh, all the cyclic currents being in the same direction. The current in any branch is thus the difference between the cyclic currents in the meshes it separates. The following problems indicate how the method may be applied.

Example.—Two liquid resistances, P and Q, of 10 and 5 ohms re-

spectively, are connected in parallel and a battery of e.m.f. 6 volts and internal resistance 4 ohms is used to send a current through them. Find the current in the two liquids if the e.m.f. of polarization in P is 0·2 volt and in Q 1·5 volt.

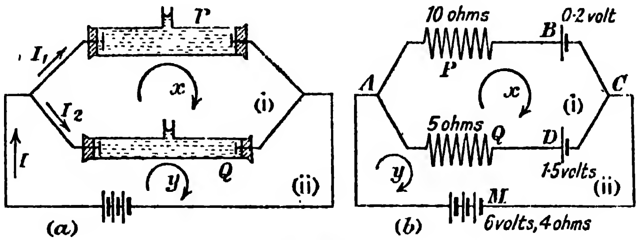


FIG. 45·10.

The circuit and a system of wire resistances and batteries, of negligible resistance but whose e.m.fs. are equal to the back e.m.fs. in the liquid resistances, are indicated in Fig. 45·10 (a) and (b) respectively. Let the cyclic currents in the mesh ABCD be x; in ADM let it be y. Then applying Kirchhoff's second law to the first mesh we have

$$10x + 5(x - y) = 1{\cdot}5 - 0{\cdot}2.$$

For the second mesh,

$$5(y - x) + 4y = 6 - 1{\cdot}5.$$

$$\therefore\ x = 0{\cdot}31 \text{ amp.};\quad y = 0{\cdot}67 \text{ amp.}$$

The current through P is equal to x; that through Q is $(y - x)$. These currents are 0·31 amp. and 0·36 amp. respectively.

Elementary Theory of the Wheatstone Bridge Network of Conductors.—Let us assume that the four resistances P, Q, R and S, arranged as in Fig. 45·11, are 'balanced,' i.e. there is no current through the gàlvanometer G when the cell B of e.m.f. E is inserted. Let G and B be the resistances of the galvanometer and battery respectively, and let x, y, and z, be the cyclic currents in the meshes (i), (ii) and (iii). Applying Kirchhoff's second law to each mesh in turn, we obtain,

$$Px + G(x - y) + R(x - z) = 0 \quad \ldots\ldots \quad \text{(i)}$$

$$Qy + S(y - z) + G(y - x) = 0 \quad \ldots\ldots \quad \text{(ii)}$$

$$R(z - x) + S(z - y) + zB = E \quad \ldots\ldots \quad \text{(iii)}$$

If $x - y = 0$, i.e. the bridge is balanced, from (i) and (ii) we obtain

$$(P + R)x - Rz = 0,$$

$$(Q + S)x - Sz = 0,$$

$$\therefore\ \frac{P + R}{R} = \frac{Q + S}{S},$$

or

$$\frac{P}{R} = \frac{Q}{S}.$$

[By solving equations (i), (ii) and (iii), $(x - y)$ not being zero, we can obtain the current through the galvanometer when the bridge is not balanced.]

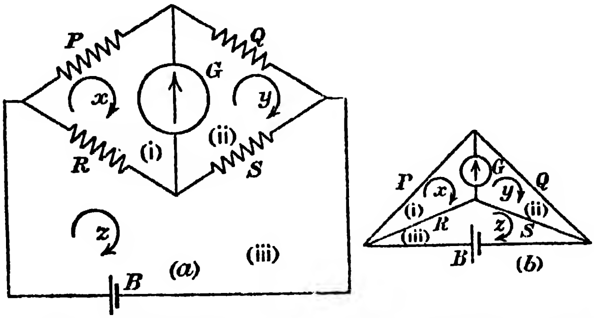


FIG. 45·11.—Elementary Theory of the Wheatstone Bridge.

Conjugate Conductors.—If two branches of any network of conductors are arranged so that an electromotive force introduced into, or existing in, one branch causes no current to flow through the other, the conductors forming those branches are termed ***conjugate conductors.*** The 'battery arm' and 'the galvanometer arm' of a Wheatstone bridge network are conjugate conductors, provided that the resistances of the other arms satisfy the usual Wheatstone bridge relationship.

Shunts.—In the construction of all sensitive galvanometers the wire on the movable bobbin has a very small diameter, a fact which limits its current carrying capacity. When it is desired to measure a large current the two terminals of the galvanometer are joined together by a short piece of thick copper wire. This allows most of the large current to pass through the thick wire, whilst only a very small current passes through the instrument. In Fig. 45·12 let G be the galvanometer and S its shunt, joined to the terminals A and B. Furthermore let these letters designate the resistances of the galvanometer and shunt. Imagine that I is the current in the main circuit, which branches at A into currents I_1 and I_2 through G and S respectively. Now the drop in potential between A and B is the same whether one goes via G or via S. In the first circuit

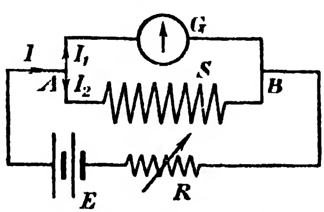


FIG. 45·12.—Use of a Shunt.

$$V_A^B = I_1G,$$

and in the second circuit

$$V_A^B = I_2 S.$$

But
$$I = I_1 + I_2 = V_A^B\left(\frac{1}{G} + \frac{1}{S}\right).$$

Hence

$$\therefore \frac{I}{I_1} = \frac{V_A^B\left(\frac{1}{G} + \frac{1}{S}\right)}{\frac{V_A^B}{G}} = \frac{\frac{S+G}{GS}}{\frac{1}{G}} = \frac{S+G}{S}.$$

This fraction measures the ratio of the current in the main circuit to that in the galvanometer. It is called the ***multiplying power*** of the shunt.

A Universal Shunt.—By using shunts having resistances $\frac{1}{9}$, $\frac{1}{99}$, and $\frac{1}{999}$, that of a galvanometer, the sensitivity of that particular galvanometer may be reduced 10, 10^2, or 10^3 times. Each galvanometer must therefore be provided with its own set of shunts, unless an AYRTON and MATHER universal shunt is available. This consists of a high resistance S, Fig. 45·13, in parallel with the galvanometer. Let us suppose that a current I enters at A and leaves at B. Then the current through the galvanometer is

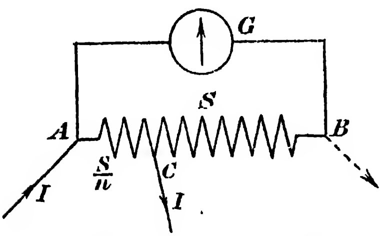


FIG. 45·13.—Principle of a Universal Shunt.

$$I\frac{S}{G+S}.$$

Let us now assume that the current leaves at C, the resistance of AC being $\frac{1}{n}$th that of S. Then a resistance S/n is in parallel with a resistance $G + \left(1 - \frac{1}{n}\right)S$. The current through the galvanometer is then

$$\frac{I\frac{S}{n}}{G + \left(1 - \frac{1}{n}\right)S + \frac{S}{n}} = \frac{1}{n}\cdot I \cdot \frac{S}{G+S},$$

i.e. the current through the galvanometer is $\frac{1}{n}$th its previous value.

In moving the point of contact from B to C, the equivalent resistance between the current leads decreases from

$$\frac{GS}{G+S} \text{ to } \frac{\frac{S}{n}\left[G + \left(1 - \frac{1}{n}\right)S\right]}{G+S},$$

so that I will, in general, be altered. For I to remain constant, the above equivalent resistances must be equal, i.e.

$$G = \frac{1}{n}\left[G + \left(1 - \frac{1}{n}\right)S\right]$$

or $$S = nG.$$

Experimental Determination of the Resistance of a Tangent Galvanometer.—(i) By a 'shunt' method. Let us assume that the resistance of the '50 turns' coil is to be determined. R, Fig. 45·14, is a resistance coil of about 40 ohms. S is a variable resistance used to shunt the galvanometer. B is a battery. These are arranged as in the diagram. It will be assumed that the resistance of G is small

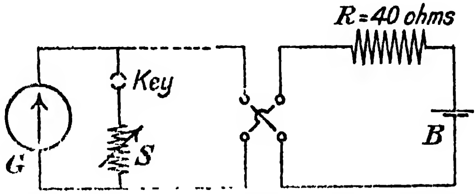


Fig. 45·14.—Shunt Method for Determining the Resistance of a Galvanometer.

compared with R so that variations in S do not affect the current from the battery.

Let α be the mean deflexion of the galvanometer when the shunt is out, i.e. $S = \infty$. Then the current I is given by

$$I = k \tan \alpha.$$

When the shunt is S, the fraction of the current passing through the galvanometer is $I_G = I.\frac{S}{G + S} = k \tan \theta$, if θ is the deflexion.

Hence $$\frac{G + S}{S} = \frac{\tan \alpha}{\tan \theta}.$$

Suppose that a series of corresponding values of S and θ are obtained. Call $\tan \theta = \frac{1}{x}$ and $S = \frac{1}{y}$. Then

$$Gy + 1 = x \tan \alpha.$$

This is a straight line whose intercept on the y-axis is $\frac{1}{G}$. G may therefore be found.

(ii) The method described above fails when the resistance of the galvanometer is considerable, for the current in the main circuit does not remain constant. The following method is desirable. The galvanometer is connected in series with a reversing key, adjustable resistance (known), R, and an accumulator. The experiment consists in obtaining corresponding values of θ, the mean deflexion of the galvanometer needle, and of R. Then if G is the resistance of the galvanometer, B that of the cell,

$$I = k \tan \theta = \frac{E}{R + G + B} = \frac{E}{R + G},$$

since for an accumulator $B \longrightarrow 0$, where k is the reduction factor for the galvanometer and E is the e.m.f. of the battery. Thus

$$R + G = \frac{E}{k}\cot\theta.$$

If, therefore, we plot $\cot\theta = x$, $R = y$, a straight line will be obtained—its intercept on the y-axis is $-G$.

[Note.—The internal resistance of a Daniell or Leclanché cell may be obtained in a similar way. A tangent galvanometer with one or two turns must be used so that $G \longrightarrow 0$. The equation is then

$$R + B = \frac{E}{k}\cot\theta.]$$

To Measure a Current by Means of a Voltmeter.—Let us suppose that a 1 ohm coil has been inserted in an electric circuit where the current is I amp. The voltage across this coil is $(1 \times I) = I$ volt. If, therefore, a voltmeter is placed in parallel with the terminals of the 1 ohm coil, the indication, in volts, of this instrument is equal to the current in amperes. This method fails if the voltmeter has a low resistance: for consider a voltmeter in which the resistance is 200 ohms. The equivalent resistance R of a 1 ohm and 200 ohm coil in parallel is given by

$$\frac{1}{R} = \frac{1}{1} + \frac{1}{200} = 1{\cdot}005, \text{ i.e. } R = 0{\cdot}995 \text{ ohm.}$$

If the current is 1 ampere in the main circuit, the voltage across the 1 ohm coil, which is recorded by the voltmeter, is $1 \times 0{\cdot}995 = 0{\cdot}995$ volts. Accordingly the indicated reading of the current is 0·995 ampere—an error of 0·5 per cent. Voltmeters generally have a resistance of at least 1000 ohms, so that the error is negligible for all practical purposes.

Moving-Coil Permanent Magnet Instruments for Measuring Steady Currents and Potential Differences.—The principle which is used to construct sensitive moving-coil galvanometers finds another application in industrial or laboratory instruments for the measurement of steady currents and potential differences. Essentially these instruments consist of a permanent magnet M, Fig. 45·15 (*a*), with soft-iron pole-pieces N, S, which have cylindrical surfaces. Between these surfaces there is mounted coaxially a cylindrical piece of soft iron, known as the core, C, and held in position by means of the brass plate B. A rectangular coil of fine copper wire, which must not contain even a trace of ferromagnetic impurity, is wound on an aluminium or copper frame, and is suspended on jewelled bearings or pivots so that it may rotate in the air gap between the core and pole-pieces. Such a method of mounting the coil permits the instrument to be carried about without risk of injury. The magnetic field in this annular air-space round the core is radial, i.e. in the gap it is directed towards the axis of C. It should be noted that this field is not of uniform

strength, but that at a constant distance from the axis of C its value is constant for an angular deflexion of about 45° on either side of the plane bisecting the pole-pieces. The current which operates the instrument is led into and out of the coil through the two phosphor-bronze hair springs shown in Fig. 45·15 (*b*) and (*c*).

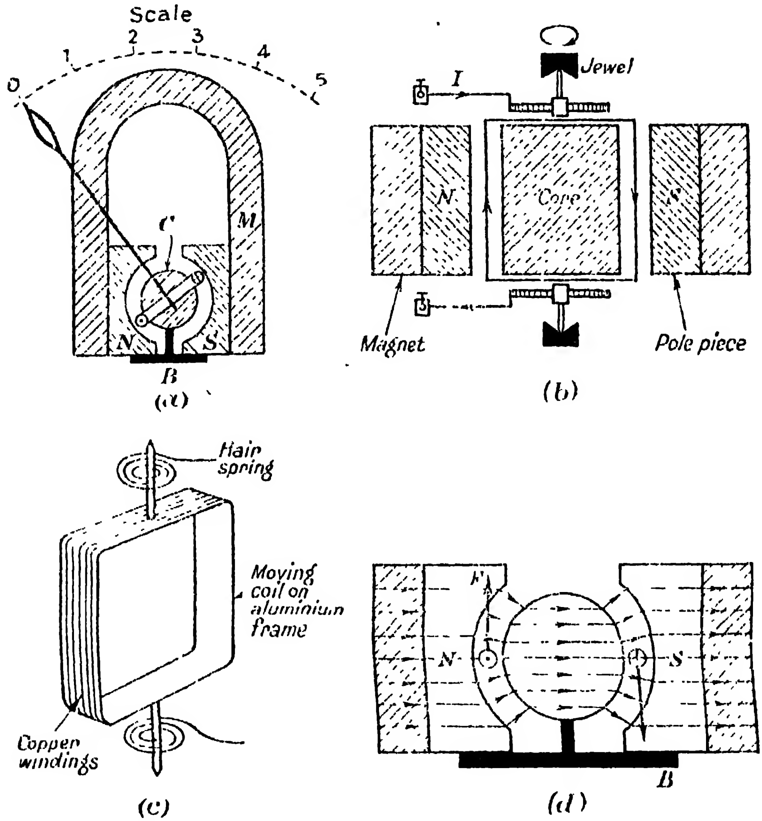


FIG. 45·15.—The Principle of a Moving-Coil Ammeter (or Voltmeter).

Now it has been shown [cf. p. 832] that the couple on the coil is proportional to the current flowing through it: the coil therefore rotates about a vertical axis until the deflecting couple is balanced by the torsional couple set up in the hair springs. A light aluminium needle rigidly attached to the moving system indicates the amount of rotation of the coil. The metal frame on which the coil is wound serves to damp the motion of the coil so that it does not oscillate about its equilibrium position: this damping effect is due to the formation of eddy currents in the former [cf. p. 950].

When such an instrument has been constructed the current required to produce a full-scale deflexion must be measured: then for whatever purpose the instrument is required this current must always pass through the coil in order to produce a full-scale deflexion. Since the magnetic field in the air-gap is strong [500 oersted, or more] the instrument is practically unaffected by stray magnetic fields: moreover, the magnet is 'aged' so that its field remains constant for long periods. Consequently one calibration suffices. Further discussion will be facilitated by the following worked examples :—

Example.—A moving-coil instrument has a resistance of 425 ohm. and a current of $1{\cdot}43 \times 10^{-4}$ amp. produces a full-scale deflexion. It is desired to use this instrument for one of the following purposes :—

(*a*) To measure a maximum current of 1 mA.
(*b*) ,, ,, ,, ,, ,, ,, 0·1 amp.
(*c*) ,, ,, ,, ,, ,, ,, 10 amp.
(*d*) ,, ,, ,, ,, potential difference of 100 mV.
(*e*) ,, ,, ,, ,, ,, ,, ,, 500 volt.

Since the current through the coil is $1{\cdot}43 \times 10^{-4}$ amp. whenever the coil is deflected to its full extent, shunts must be provided if larger currents are to be measured. Thus in (*a*), let S, Fig. 45·16 (*a*), be the shunt: the current I, through it, is $(1 \times 10^{-3}) - (1{\cdot}43 \times 10^{-4})$ amp. when the coil is fully deflected. Now the p.d. across the instrument is equal to the p.d. across the shunt. This fact provides the key to the problem, for we have

$$\begin{aligned} 1{\cdot}43 \times 10^{-4} \times 425 \text{ volt.} &= \text{p.d. across moving coil} \\ &= \text{p.d. across S} \\ &= (10 \times 10^{-4} - 1{\cdot}43 \times 10^{-4}) \times \text{S} \end{aligned}$$

$$\therefore \text{S} = \frac{1{\cdot}43 \times 425}{8{\cdot}57} = 70{\cdot}8 \text{ ohm.}$$

In (*b*) the current through the shunt is $(1000 - 1{\cdot}43)10^{-4}$ amp. Hence, proceeding as before,

$$1{\cdot}43 \times 10^{-4} \times 425 = (1000 - 1{\cdot}43)10^{-4}\text{S}$$

$$\therefore \text{S} = \frac{1{\cdot}43 \times 425}{999{\cdot}57} \doteq \frac{1{\cdot}43 \times 425}{1000} = 0{\cdot}608 \text{ ohm.}$$

In (*c*) the current through the shunt may be taken as 10 amp.

$$\therefore \text{S} = \frac{1{\cdot}43 \times 10^{-4} \times 425}{10} = 0{\cdot}0061 \text{ ohm.}$$

When the instrument is to be used to measure potential differences a resistance R must be placed in series with the moving-coil instrument, the whole being placed in parallel with the piece of apparatus across which it is desired to measure the potential difference.

Thus in (*d*) the current through both R and the moving-coil instrument is $1{\cdot}43 \times 10^{-4}$ amp., while the p.d. across R and G together is 0·1 volt. Hence

$$0{\cdot}1 = 1{\cdot}43 \times 10^{-4}[425 + \text{R}]$$

$$\therefore \text{R} = \frac{10^3}{1{\cdot}43} - 425 = 699 - 425 = 274 \text{ ohm.}$$

Again, in (*e*)

$$500 = 1{\cdot}43 \times 10^{-4}[425 + R].$$

$$\therefore R = 3{\cdot}5 \times 10^{6} \text{ ohm.} = 3{\cdot}5 \text{ megohm.}$$

This value of the resistance is so large that a less sensitive moving-coil instrument would be made: the added resistance would then be smaller. In precision voltmeters 'a thousand ohms per volt to be measured' is the standard usually set for the value of the added resistor or 'multiplier.'

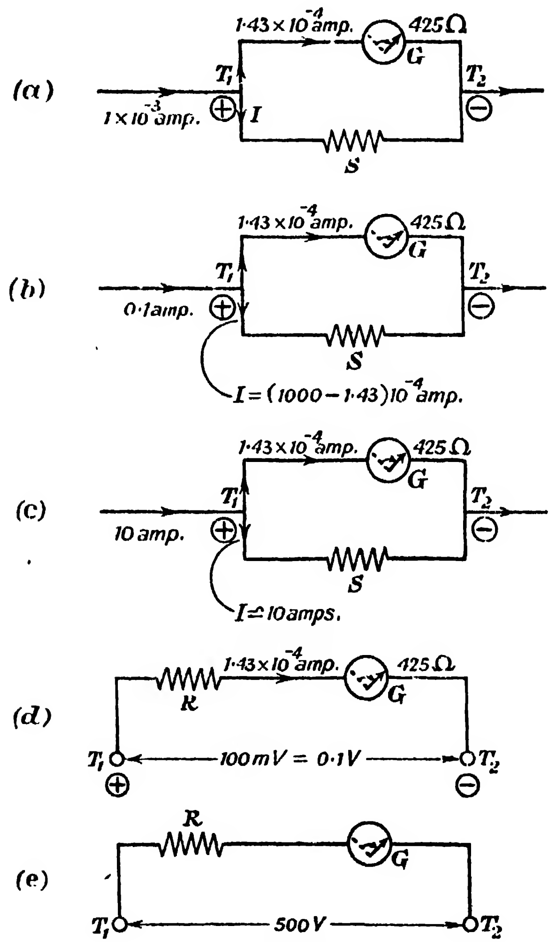


FIG. 45·16.

Note on the Adjustment of an Ammeter or a Voltmeter.—When the resistance of the shunt required to convert a given moving-coil instrument into an ammeter has been calculated a piece of manganin, cut from a thick wire or sheet, is selected, its

resistance being preferably slightly below the calculated value. Short copper wires P_1 and P_2 are hard-soldered on in the positions indicated in Fig. 45·17 (*a*) and the galvanometer G connected permanently in position. A current measured on a standard instrument is then passed between the external terminals T_1 and T_2 of the ammeter and, if necessary, the resistance of the shunt increased by scraping away a portion of the metal until the indications of the two instruments are identical at the full-scale value of the ammeter under construction. If the ammeter reading is in excess of that of the standard instrument the resistance of the shunt S is too high and a high resistance shunt must be placed across T_1T_2 and its value adjusted until the discrepancy disappears.

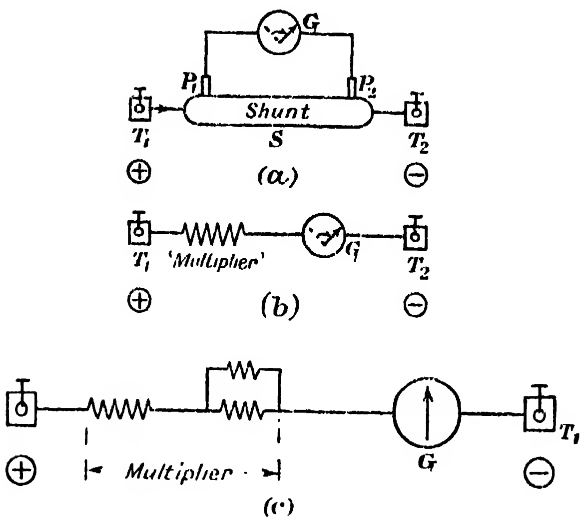


Fig. 45·17.

When voltmeters are being adjusted the resistance of the multiplier, Fig. 45·17 (*b*), can generally be measured sufficiently accurately and adjusted before it is inserted in the instrument. If not, it could be made in excess of the required value and then adjusted by placing a shunt across a portion of the multiplier—as in Fig. 45·17 (*c*). It must be noted that when a voltmeter is being compared with a standard voltmeter, the two instruments are arranged in parallel.

If it is desired to convert an ammeter into a voltmeter, the shunt must be removed and the appropriate resistor placed in series with the galvanometer portion of the instrument. In the reverse case, the resistor is removed and the necessary shunt placed in position.

Kelvin's Method for Determining the Resistance of a Galvanometer.—The galvanometer is placed in the fourth arm of a post-office box and a tapping key across BD, the position usually occupied by the galvanometer—cf. Fig. 45·18. Since under these conditions the current through it would be excessive, only a small potential difference is applied across CA [cf. p. 848]. P, Q, and R are then so arranged that the deflexion of the galvanometer is of convenient magnitude. The key K is then closed and, in general, there will be a change in this deflexion. R is changed until there is no change in the deflexion when K is opened or closed. When this occurs the points B and D must be at the same potential, and we have $\frac{P}{Q} = \frac{R}{X}$, where X is the galvanometer resistance. In carrying out this experiment the current through the bridge should be reversed and the observations repeated.

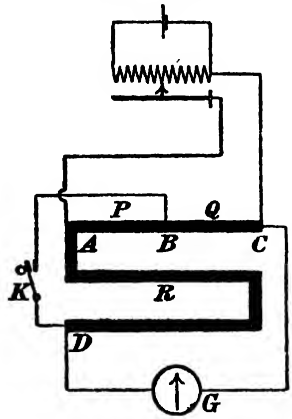


FIG. 45·18.—Resistance of a Galvanometer. Kelvin's Method.

Mance's Method for Determining the Internal Resistance of a Cell.—The cell is placed in the fourth arm of a post-office box; then with the galvanometer, G, and a high resistance, Z, across BD,

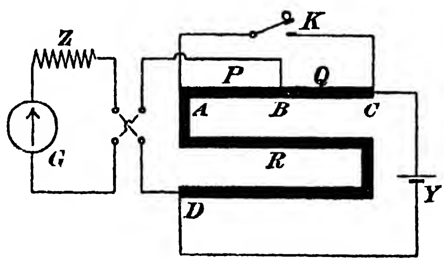


FIG. 45·19.—Internal Resistance of a Cell. Mance's Method.

a tapping key is placed across AC—cf. Fig. 45·19. The high resistance, Z, is necessary to limit the current through G. The experiment consists in adjusting P, Q, and R so that there is no change in the galvanometer deflexion when K is opened or closed. Then $\frac{P}{Q} = \frac{R}{Y}$, where Y is the resistance of the cell. The proof of this statement may be found in a text-book of Practical Physics.

The 'End Corrections' of a Metre Bridge.—The small resistances of contact at the ends of a metre bridge wire and errors arising from the fact that the wire may not be exactly 100 cm. long may be determined as follows:—Resistances of 1Ω and 101Ω are placed in the two gaps of the bridge ABCD—Fig. 45·20 (*a*)—and a balance point D on the wire is located in the usual way. If l and $(100 - l)$

are the lengths into which D divides the wire, and α_1 and α_2, expressed as cm. of bridge wire, the 'end corrections' we are endeavouring to find, the usual Wheatstone bridge relation gives

$$\frac{1}{101} = \frac{l + \alpha_1}{(100 - l) + \alpha_2}.$$

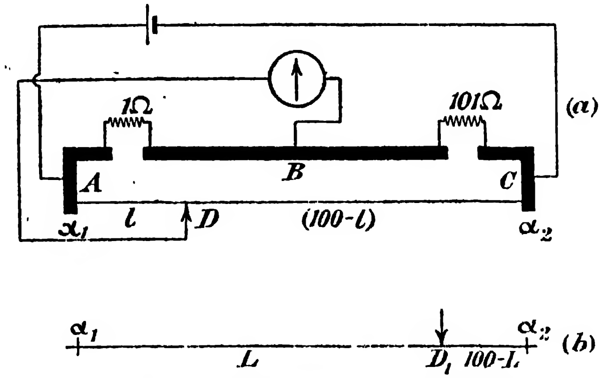


FIG. 45·20.—End Corrections to a Bridge Wire.

The coils are then interchanged and a new balance point D_1 found. Referring to Fig. 45·20 (*b*) where this is indicated, we have

$$101 = \frac{L + \alpha_1}{100 - L + \alpha_2}.$$

Writing these equations in the form

$$101(\alpha_1 + l) = \alpha_2 + 100 - l,$$

and

$$\frac{L + \alpha_1}{101} = 100 - L + \alpha_2,$$

we have, by subtraction,

$$101(\alpha_1 + l) - \frac{L + \alpha_1}{101} = L - l.$$

$$\therefore \alpha_1 = \frac{L - 101l}{100}.$$

Similarly
$$\alpha_2 = \frac{101L - l - 100^2}{100}.$$

A graphical method for the same purpose has been suggested by FERGUSON. Suppose R_1 and R_2 are the known resistances placed in the 'gaps' of the bridge. Then if l_1 and l_2 are the lengths into which the bridge wire is divided when the bridge is balanced

$$\frac{l_1 + \alpha_1}{l_2 + \alpha_2} = \frac{R_1}{R_2} = \gamma \text{ (say)}.$$

This may be written

$$\gamma l_2 - l_1 = -\gamma\alpha_2 + \alpha_1.$$

The bridge is balanced for a series of values of $\frac{R_1}{R_2}$, i.e. of γ; then if we call

$$\gamma l_2 - l_1 = y; \quad \gamma = x,$$

we have
$$y = -\alpha_2 x + \alpha_1.$$

Hence α_1 and α_2 may be derived from the intercept on the y-axis and the slope of the straight-line graph.

Resistivity.—By means of a metre bridge or post-office box it may be shown that the resistance, R, of a uniform wire is directly proportional to its length, l, and inversely proportional to its cross-sectional area, a, provided, of course, that the physical state of the wire remains constant. Hence $R = \frac{\chi l}{a}$, where χ is a constant for the material of the wire. It is termed its ***resistivity***, or ***specific resistance.***

Since we may write $[\chi] = \left[\frac{Ra}{l}\right]$ = [ohm. × length² ÷ length],

the unit of resistivity is the ohm.cm. when the unit of length is the centimetre. It should be noted that χ is ***not*** equal to the resistance of a unit cube, for this is measured in ohms, whereas the resistivity is measured in ohm.cm. The two quantities are numerically equal provided that the lines of flow of the current are parallel to one edge and therefore to four edges of the cube. This latter limitation is necessary, for if the current entered at one corner and left the cube at the diagonally opposite corner it is difficult to calculate the resistance offered by the cube to the current.

Conductivity.—The ***conductivity***, or ***specific conductance*** of a substance is the reciprocal of its resistivity. It is therefore expressed in ohm.$^{-1}$ cm.$^{-1}$ and is denoted by σ.

EXAMPLES XLV

1.—Calculate the p.d. across a lamp whose resistance is 104 ohm. if the current is 1·06 ampere.

2.—A current from a battery passes through 10 ohm. and a tangent galvanometer. The reduction factor (k), [where $I = k \tan \theta$] for the galvanometer is 0·63 amp. The deflexion observed is 47°. If the resistances of the battery and galvanometer are negligible, calculate the e.m.f. of the battery.

3.—A battery consists of 3 cells arranged in parallel. Each cell has an e.m.f. 1·08 volt., and a resistance 3·5 ohm. What current will the battery send through a 10 ohm. resistance ?

4.—A cell whose internal resistance is 0·52 ohm produces a current of 0·27 ampere in a 6-ohm. wire. Find the e.m.f. of the cell, and the difference in potential which exists between its terminals. [This p.d. is equal to the p.d. across the 6 ohm. coil.]

5.—A battery is connected in series with a tangent galvanometer of resistance 18 ohm., and the deflexion observed is 54°. When an

additional resistance of 12 ohm. is placed in the circuit the deflexion is reduced to 42°. What is the battery resistance ?

6.—Calculate the resistance of the following coils when arranged (*a*) in series, (*b*) in parallel—2 ohm., 3 ohm., 4 ohm. What is the current through a cell whose e.m.f. is 2·08 volt. when this is connected in turn to each arrangement ?

7.—Define the resistivity of a substance. A wire has a resistance of 40 ohm. It is cut in halves and the two portions arranged in parallel. What is the resistance of the combination ?

8.—A coil has a resistance of 20·37 ohm. What must be the value of the shunt resistance so that the whole may be equivalent to a 20 ohm. coil ?

9.—A galvanometer has a resistance of 1064 ohm. What must be the shunt so that only one-tenth of the current shall pass through the galvanometer ?

10.—ABCD is a square, each side of which has a resistance of 2 ohm. A 5 ohm. coil is placed across AC. Calculate the equivalent resistance between A and C, and also between B and D.

11.—A coil having 8 turns of wire, each 1 metre in diameter, is placed with its plane in the magnetic meridian. Calculate the value of H_0 if a current of 1·6 ampere. deflects the needle through 45°.

12.—Two cells are placed in series with a tangent galvanometer and a resistance. The deflexion is 50° when the cells assist one another, whilst it is only 22° when they are in opposition. Calculate the e.m.f. of the larger cell if that of the smaller is 1·08 volt.

13.—A and B are two points on the circumference of a circle consisting of uniform wire. They subtend an angle of 127° at the centre. If A and B are joined to a battery, calculate the ratio of the currents in the two segments of the wire.

14.—State Ohm's law and describe how you would verify it for a conductor in the form of a long thin wire. A, B, C, and D are four coils of wire of 2, 2, 2, and 3 ohm. resistance respectively, arranged to form a Wheatstone bridge network. Calculate the value of the resistance with which the coil D must be shunted in order that the bridge may be balanced. If the shunt is a wire 100 cm. long and 0·2 mm. diameter, calculate the resistivity of the material.

15.—Establish the relation between the resistances of the arms of a balanced Wheatstone's bridge. How may the ordinary arrangements of a Wheatstone bridge be modified for finding the resistance of the electric cell used ?

16.—If the wire of a Wheatstone bridge has a resistance of 1 ohm and the bridge is used to compare the resistances of 2 ohm. and 3 ohm. respectively, what current flows along the wire when the galvanometer shows no deflexion if the battery used has an e.m.f. of 1·7 volt. and an internal resistance of 5 ohm. ? (L. '28.)

17.—Establish a formula for calculating the equivalent resistance of two conductors joined in parallel. The terminals of a battery of e.m.f. 10 volt. and of negligible internal resistance are connected to two coils each of 100 ohm. resistance, joined in series. A voltmeter of resistance 500 ohm. is connected in turn across (*a*) each of the coils, (*b*) the terminals of the battery. What is the reading of the instrument in each instance ?

18.—Describe the construction and explain the action of a moving-coil voltmeter. A certain voltmeter has a range of 15 volt. and a

resistance of 1000 ohm. How would you use it to measure voltages up to 150 volt. ?

19.—Describe some form of sensitive galvanometer. A galvanometer of 100 ohm. resistance gives a full-scale deflexion for a current of one-tenth of a milliampere. How would you arrange so that it could be used as a voltmeter giving a full-scale deflexion for 1 volt. ?

20.—Explain the action of a shunt. A current from a battery of resistance 4 ohm. is sent through an electric heater of resistance 10 ohm. With what resistance must the heater be shunted in order to decrease the amount of heat developed in it to half its former value ?

21.—Explain the theory and the method of using a potentiometer (*a*) to compare the electromotive forces of two cells, (*b*) to calibrate an ammeter.

22.—P, Q, R, S, are resistances taken in cyclic order in a W.B. network. P and Q are the ratio coils: S is the unknown resistance and R a 20 ohm coil which needs to be shunted with 350 ohm. to secure an exact balance. When P and Q are interchanged balance is restored by altering the shunt across R to 498 ohm. Find the resistance of S and the ratio P : Q. (L.I.)

23.—Explain how a 'shunt' may be used to alter the sensitiveness of a galvanometer. A voltmeter reading from 0 to 10 volt. has a resistance of 1000 ohm. How would you convert it into an ammeter with a range from 0 to 1 ampere ?

24.—'In all direct current galvanometers there is called into play a force of automatically varying moment which serves to balance the electromagnetic moment due to the current, and to restore the recording needle to the zero position when the current is switched off.' Explain this statement by means of descriptions of various types of galvanometer. (N.H.S.C. 29.)

25.—A metal tube of length l has internal and external radii a and b respectively. If χ is the resistivity of the material of the tube, show that R, the resistance of length l of the tube is given by

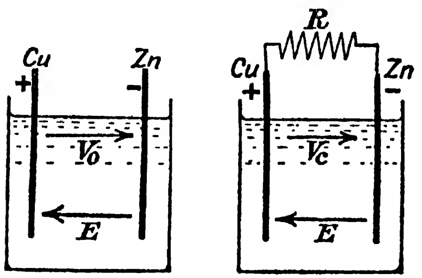
$$R = \frac{\pi}{\chi l}(b^2 - a^2).$$

Need the axes of the cylindrical surfaces be coaxial ?

CHAPTER XLVI

ELECTROMOTIVE FORCE, THE POTENTIOMETER, AND SOME ELECTRICAL MEASUREMENTS

Electromotive Force and the Internal Resistance of a Cell.—At the beginning of Chapter XLIII it was shown that the potential difference between a copper and a zinc electrode placed in dilute sulphuric acid was caused by an electromotive force in the cell. A similar statement is true for all cells but we shall consider the simple cell as a concrete example. When the electrodes are not connected by a wire and steady conditions have been reached [almost instantaneously] the potential difference between the electrodes is numerically equal to the electromotive force of the cell. The potential difference between the plates tends to make electricity pass from the copper to the zinc along a path not in the cell, whereas the electromotive force is only operative inside the cell and acts from the zinc to the copper. If E is the electromotive force and V_0 the potential difference between the plates when they are not joined together, i.e. the cell is on 'open circuit,' then $E = V_0$—cf. Fig. 46·1 (*a*).



(*a*) (*b*)
FIG. 46·1.

When the plates are connected by a wire of resistance R electricity immediately begins to flow in a direction from the copper to the zinc along the wire. This is caused by the potential difference across the wire. Inside the cell the electromotive force is still operative for chemical reactions are taking place in it. Opposing this e.m.f. there is V_c the potential difference between the plates. This is often termed the electromotive force of the cell on 'closed circuit,' but this is really a misnomer, for the electromotive force of the cell is constant and it is only the potential difference between its plates which varies with the current supplied by the cell. Since $E > V_c$ electricity will be driven through the cell from zinc to copper and the current in the cell will be $I_1 = \dfrac{(E - V_c)}{B}$, where B is the internal

resistance of the cell. Outside it the current will be $I_2 = \frac{V_e}{R}$. These two currents will differ for a small fraction of a second, i.e. until the rate at which electricity is passing from the copper electrode is equal to the rate at which it is gaining electricity. Then $I_1 = I_2$ and

$$\frac{E - V_e}{B} = \frac{V_e}{R}.$$

From the above we see that we cannot measure the electromotive force of a cell directly but must measure the potential difference between its terminals on open circuit. This may be done by applying a potential difference to the cell so that it opposes the potential difference between its plates and adjusting it so that the current from or to the cell is zero. The electromotive forces of two cells may be compared by the following methods.

The Comparison of Electromotive Forces

The Sum and Difference Method.—In order to compare the e.m.fs. of two cells they are connected in series with a tangent galvanometer and a resistance which is adjusted so that the deflexion of the needle is in the neighbourhood of 45°. The two cells are then connected so that they are in opposition, and the resistance still being as before the deflexion is again noted. Let E_1 and E_2 be the e.m.fs. of the cells; let B, G, and R be the ohmic resistances of the battery, galvanometer, and resistance box respectively, whilst θ_1 and θ_2 are the deflexions of the galvanometer. If I_1 and I_2 are the respective currents in the two circuits, then

$$I_1 = \frac{E_1 + E_2}{B + G + R} = k \tan \theta_1,$$

and

$$I_2 = \frac{E_1 - E_2}{B + G + R} = k \tan \theta_2,$$

where k is the reduction factor of the instrument. By division, and use of the lemma * below, we have

$$\frac{E_1 + E_2}{E_1 - E_2} = \frac{\tan \theta_1}{\tan \theta_2}.$$

$$\therefore \frac{E_1}{E_2} = \frac{\tan \theta_1 + \tan \theta_2}{\tan \theta_1 - \tan \theta_2}.$$

---

* *Lemma:* If
$$\frac{a}{b} = \frac{c}{d} = k \text{ (say)}$$
then
$$\frac{a + c}{a - c} = \frac{b + d}{b - d}.$$
Since $a = bk$ and $c = dk$, we have by substitution
$$\frac{a + c}{a - c} = \frac{bk + dk}{bk - dk} = \frac{b + d}{b - d},$$
i.e. if two fractions are equal, one can add and subtract the numerator and denominator of each to form fractions which are still equal. This has been done in order to solve the above equations.

The Potentiometer.—The above method of comparing voltages can only be considered as a very approximate one, mainly because the result depends upon the difference of two quantities which are of the same order of magnitude. The potentiometer is an instrument designed for the accurate comparison of potential differences. The theory of the instrument can be gathered from a consideration of Fig. 46·2. AB is a manganin wire of about 8 ohm. resistance and stretched over a scale graduated in cm., etc. It is connected to a source of constant potential C, and a key K. For convenience an accumulator is generally used at C, but it must be understood quite clearly that any appliance capable of yielding a constant voltage would do. The cells whose voltages are to be compared must not be compared with C for reasons which will be stated later. The key which is used in a potentiometer experiment should never be one of the 'plug-in' variety, since the resistance of such a plug is variable. It is much better to construct two holes in a block of wax and fill them with mercury. The circuit wires dip, one into each cup of mercury, and the distance between the cups can be bridged with a short piece of thick copper wire, the surface of which has been previously amalgamated with mercury.*

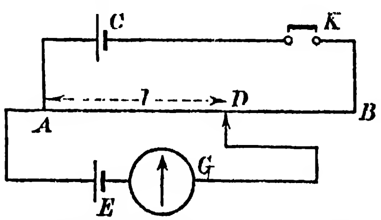


FIG. 46·2.—A Simple Potentiometer.

One of the cells to be compared is placed at E, connected to a galvanometer G, to the manganin wire at A, and finally to a sliding contact at D. The cells C and E must always be so arranged that like electrodes are joined to A. Assuming that the potential drop across AB is not less than that across E, some point D on the wire AB will have the same potential as the negative electrode of the cell at E. This point is found by moving the sliding contact along AB until the galvanometer reading is zero. Under these conditions there is no fall in potential across G and the connecting wires, for the drop in potential is equal to the product of the resistance and current, and although the resistance may be large the current is zero. It therefore follows that the potential between A and D is equal to that across E. It cannot be said that the potential of A is equal to that of the positive plate of C, because there is a current in the connecting wire and hence there must be a difference of potential between C and A.

If AB is a uniform wire the fall in potential along AD is pro-

* This is very easily done by cleaning the copper with nitric acid, and then plunging it into dilute sulphuric acid and mercury for a few seconds. The amalgamated copper is then washed with distilled water and dried.

portional to the length AD. Call this length l. Then $E = \kappa l$, where κ is a constant. When cells E_1, E_2, etc., are placed at E and the corresponding lengths, l_1, l_2, etc., determined, it follows that

$$\frac{E_1}{E_2} = \frac{l_1}{l_2}.$$

If E_2 is a standard Weston Cell, then the voltage of E_1 can be calculated. It will be observed that there is no reference at all to the cell C in this equation.

The Potentiometer in Practice.—In actual laboratory practice it is advisable to insert an adjustable resistance R in series with the potentiometer wire AB, Fig. 46·3. In addition a resistance of about 1000 Ω resistance should be placed in series with the cell E_1 under investigation. Let AD $= l_1$ where D is such that the p.d. across AD is equal to the e.m.f. of E_1. Then $E_1 = \kappa l_1$, as before. When E_1 is replaced by a second cell E_2, we shall have $E_2 = \kappa l_2$. Now it may happen that l_1 and l_2 are relatively short in comparison with AB. The experiment is repeated, R being adjusted so that the

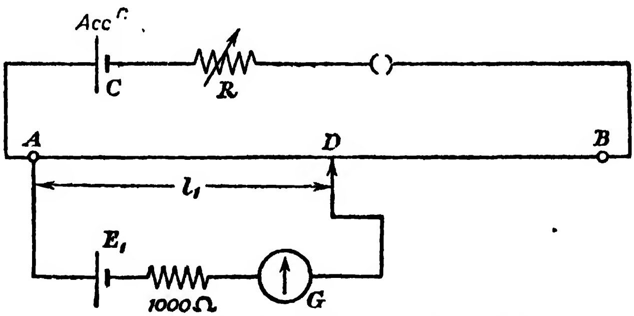


FIG. 46·3.—A Simple Potentiometer in Practice.

position of D when the cell of higher e.m.f. is in use is close to B. Then the ratio $l_1 : l_2$ may be determined with greater accuracy.

The advantage of placing the 1000 Ω resistance in series with the cell E_1 (or E_2) is that it prevents large currents from being taken from the cell while the position of D is being determined. This is very desirable when the e.m.f. of a cell does not recover when a large current has been taken from it.

Similar features may be added to the experimental arrangements which follow.

To Measure a Current Accurately.—The potentiometer is easily applied when one wishes to measure a current accurately. For this purpose the connections are arranged as in Fig. 46·4. The wire AB, the source of constant current C, and key K are as before. Suppose it is desired to measure the current in the circuit EPR,

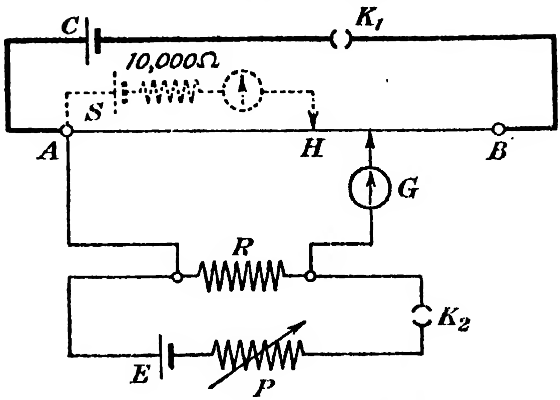


FIG. 46·4.—Use of a Potentiometer to Measure a Current.

where E is a cell sending a current through a resistance P and a standard resistance coil R when the key K_2 is closed. The terminals of the standard coil R are joined to A and to D through the galvanometer G. When there is no current in the galvanometer the voltage across R is proportional to the length AD. But the voltage (V) across a coil of resistance R is given by V = IR, where I is the current. Hence, if AD = l, IR = V = κl, where κ is a constant: it is the drop in potential per unit distance along AB. In order to obtain the value of I it is necessary to know κ. For this purpose a standard cell S is connected to a high resistance (10,000 Ω at least), a galvanometer and jockey and arranged as shown. When the deflexion in the galvanometer is zero, i.e. the jockey is at H, say, the p.d. across AH is equal to the e.m.f. of S [1·0184 volt., if S is a Weston cell]. κ is therefore known and I may be calculated. The resistance R should be such that IR is nearly one volt.: then AD and AH are nearly equal.

Experiment.—**Place a tangent galvanometer, 1 ohm. coil, adjustable resistance, key, and battery in series. Obtain a deflexion of the galvanometer needle of about 45° and measure the potential difference across the 1 ohm. coil by means of a potentiometer. Standardize the potentiometer by using a Daniell cell or other cell of known e.m.f. Assume H_0 = 0·18 oersted, and calculate the ratio of the e.m.u. of current to the practical unit of current.**

To Compare Two Resistances.—Since the potential difference across a resistance is proportional to the product of its resistance and the current through it, the ratio of two resistances will be equal to the ratio of the potential differences across them when they are each carrying the same current. Hence, by comparing these potentials we have a means of comparing two resistances.

To carry out this comparison of p.ds. a potentiometer is set up

as in Fig. 46·5 (*a*) and also a second circuit comprising the resistances R_1 and R_2, to be compared, an accumulator, adjustable resistance and key K_2, as in Fig. 46·4 (*b*). When K_2 is closed the p.ds. across R_1 and R_2 will be IR_1 and IR_2 respectively, where I is the current. The points A and P are now connected electrically while Q is connected through a galvanometer, G, to a jockey which is moved along the potentiometer wire AB, until a point D_1 is found such

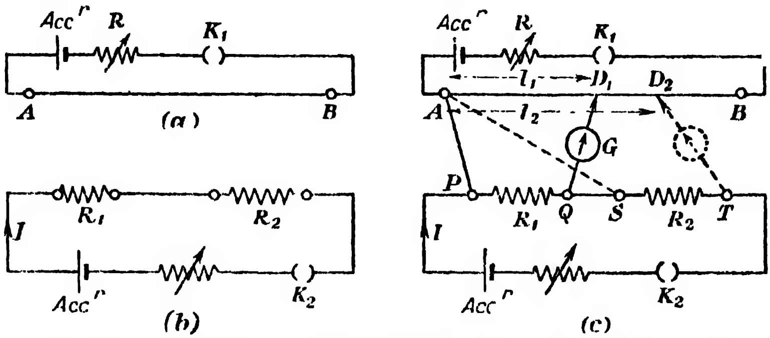


FIG. 46·5.—Comparison of Resistances by Means of a Potentiometer.

that there is no current through G. Then, with the usual notation,

$$V_1 = IR_1 = \kappa l_1.$$

The p.d. across R_2 is found by making the connections shown by the dotted line. Then

$$V_2 = IR_2 = \kappa l_2,$$

so that

$$\frac{R_1}{R_2} = \frac{l_1}{l_2}$$

If l_1 and l_2 are small the resistance R must be increased to diminish the potential drop per unit length of the potentiometer wire and the experiment repeated. The method works equally well whether the resistances are large [10^5 ohm.] or small [10^{-2} ohm.] provided that a suitable galvanometer is selected and the current adjusted accordingly.

Example.—The resistance of a potentiometer wire 100 cm. long is 5·12 Ω. What resistance, R, must be placed in series with the wire in order that the drop in potential down the wire shall be 1 mV. cm.$^{-1}$ if an accumulator of e.m.f. 2·08 V. supplies the current ?

The current, I, through R must be the same as that through the potentiometer wire. The p.d. across the wire = 1 mV. cm.$^{-1}$ × 100 cm. = 0·1 V.

$$\therefore\ I = \frac{0{\cdot}1\,V}{5{\cdot}12\,\Omega} = 0{\cdot}01953\,A.$$

$$= \frac{2{\cdot}08\,V}{(R + 5{\cdot}12)\,\Omega}.$$

$$\therefore\ R = 101{\cdot}4\,\Omega.$$

This resistance is so large that the fact that we have assumed the initial resistance of the accumulator to be zero, is of no practical importance.

The Rayleigh Potentiometer.—Two identical resistance boxes AB and CD, Fig. 46·6, are connected in series. The 'variable arms' of two post-office boxes may be used. The plugs are all removed from AB whilst none is removed from CD. To compare the e.m.fs. of two cells, E_1 and E_2, they are arranged so that either may be connected through a galvanometer G to the ends of AB. If E_1 has thus been connected, plugs are inserted in AB and the corresponding plugs removed from CD until the galvanometer deflexion is zero. The p.d. across AB is then equal to the e.m.f.

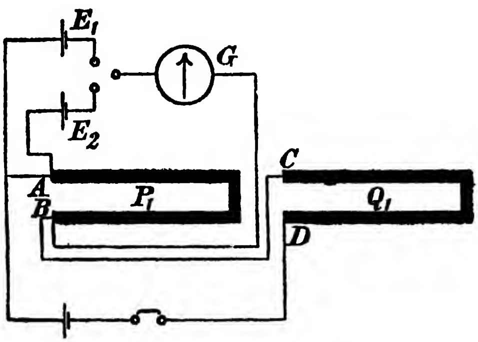


FIG. 46·6.—A Rayleigh Potentiometer.

of E_1. Let P_1 be the value of resistance of the unplugged coils in AB. The experiment is repeated with E_2 and the corresponding value P_2 found. Since the total resistance in AB and CD has been kept constant, the current through them has been kept constant, so that $\frac{E_1}{E_2}=\frac{P_1}{P_2}$. Generally, the total resistance is 11,000 ohm. With a sensitive galvanometer changes produced by transferring one ohm from AB to CD may be detected by this method, so that e.m.fs. may be compared with an error of less than 0·01 per cent.

Internal Resistance of a Cell.—At the beginning of this chapter we proved that the potential difference between the plates of a cell depended upon its internal resistance and the current it was supplying. Let I be the current supplied by the cell through a known resistance R and let E be the electromotive force of the cell. Then if V_e is the p.d. between the plates, we have

$$I=\frac{V_e}{R}=\frac{E-V_e}{B}.$$

If E and V_e are measured by a voltmeter, B may be calculated since the above equation may be written

$$B = \left(\frac{E - V_e}{V_e}\right) . R . = \left(\frac{E}{V_e} - 1\right)R.$$

Instead of using a method involving the direct measurement of E and V_e, we may compare E and V_e with the aid of a potentiometer. To do this the apparatus is arranged as in Fig. 46·7. First, the resistance R_1 should be adjusted so that the drop in potential along

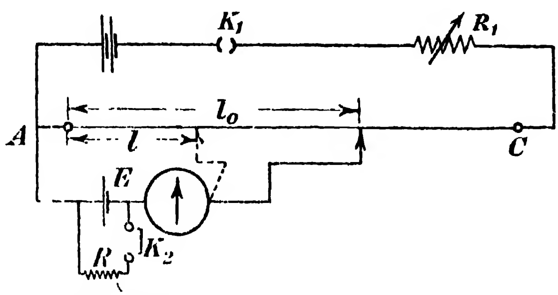


FIG. 46·7.—Internal Resistance of a Cell by a Potentiometer Method.

the potentiometer wire AC is just greater than the e.m.f. of the cell E, whose internal resistance, B, is required. Then by inserting the key K_2 the resistance R is placed across the cell. Let l be the potentiometer reading corresponding to the p.d. between the plates under these conditions. The current supplied is $\frac{E}{R + B}$, so that the p.d. across R [which is the so-called e.m.f. of the cell on closed circuit] is

$$\frac{E}{R + B} . R.$$

This is equal to kl, where k is a constant. Thus

$$E . \frac{R}{R + B} = kl,$$

which may be written

$$\frac{E}{k} . \left(\frac{R}{l}\right) = R + B.$$

A series of readings with different values of R should be taken, when the following graphical method may be used to find B.

Calling $\frac{R}{l} = x$, and $R = y$, the above equation becomes

$$y = - B + \frac{E}{k}(x)$$

The experimental points should therefore lie on a straight line whose intercept on the y-axis is $-$ B.

The Carey Foster Bridge for Comparing Nearly Equal Resistances.—The difference between the resistances of two coils whose resistances are nearly equal may be accurately determined by a modified form of the Metre Bridge due to CAREY FOSTER. A uniform wire AB, Fig. 46·8, is stretched across a scale in cm., etc., and the two resistances to be compared, R and S, are placed in the outer gaps of a copper bar whose extremities are joined to AB. P and Q are two resistances placed in the inner gaps of this bar. These should

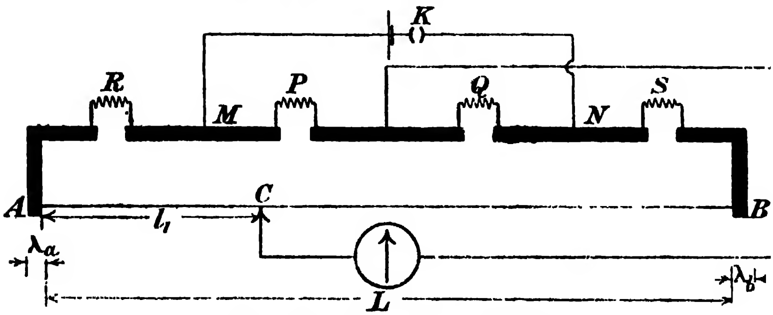


FIG. 46·8.—Carey Foster Bridge.

be nearly equal, but it is not necessary to know what are their actual resistances. A battery, reversing key, K, and galvanometer are connected as indicated. If the difference in resistance between R and S is not large a point C on the wire AB may be found where the galvanometer deflexion is zero. Let AC $= l_1$, AB $=$ L, the total length of the wire, while λ_a and λ_b are the end corrections expressed as lengths of bridge wire. Then P and Q may be regarded as two arms of a Wheatstone bridge network, while R and S and the resistances of the circuit between them and C constitute the other two arms. When the bridge is balanced we have

$$\frac{P}{Q} = \frac{R + \rho(\lambda_a + l_1)}{S + \rho(\lambda_b + L - l_1)}$$

where ρ is the resistance per cm. of the wire.

The coils R and S are now interchanged and a new balance point on the bridge wire determined. Let this be at distance l_2

Then

$$\frac{P}{Q} = \frac{S + \rho(\lambda_a + l_2)}{R + \rho(\lambda_b + L - l_2)}.$$

Hence

$$\frac{R + \rho(\lambda_a + l_1)}{S + \rho(\lambda_b + L - l_1)} = \frac{S + \rho(\lambda_a + l_2)}{R + \rho(\lambda_b + L - l_2)}.$$

$$\therefore \frac{R + \rho(\lambda_a + l_1)}{R + S + \rho(\lambda_a + \lambda_b + L)} = \frac{S + \rho(\lambda_a + l_2)}{S + R + \rho(\lambda_a + \lambda_b + L)}.$$

Since the denominators of these fractions are equal,

$$R + \rho(\lambda_a + l_1) = S + \rho(\lambda_a + l_2)$$

or

$$R - S = \rho(l_2 - l_1).$$

This equation expresses the difference between R and S in terms of the resistance of the wire and we note that the end corrections λ_a and λ_b do not appear in it.

To determine ρ, R is replaced by a 1-ohm. coil and S by a 1-ohm. coil shunted by a 10-ohm. coil, i.e. S = 0·909 ohm. If x_1 and x_2 are the bridge readings when these are used in the above manner, we have

$$1 - 0{\cdot}909 = \rho(x_2 - x_1)$$

so that ρ is known.

To obtain accurate results the coils must be connected to the terminals with the aid of thick copper strips, the contacts well cleaned and the battery connections reversed. By taking the mean of the readings for each arrangement errors due to parasitic e.m.fs. (thermo-electric, etc.) will be eliminated.

Determination of a Small Resistance in Terms of Standard Resistance Coils.—Let AB, BC, and CA, Fig. 46·9, be coils having resistances 1 ohm., 1 ohm. and 10^4 ohm. respectively. HK is, for example, a copper rod and we desire to determine the resistance of this rod between the potential leads connected to points M and N on it. A and B, and M and N, are connected to mercury cups, *c, d, e, f*, whilst a high resistance galvanometer is connected to two other such cups, *a*, *b*. Let I be the current through the rod when the circuit is connected to the battery S. Since the resistance of AC is very large, I will also be the value of the current through CB, and the current through AC and AB will be I × 10^{-4}. The potential difference across AB is therefore I × 10^{-4}, whilst that across MN is I × r where r is its resistance. If θ_1 and θ_2 are the readings of the high resistance galvanometer, G, when across AB and MN respectively, we have

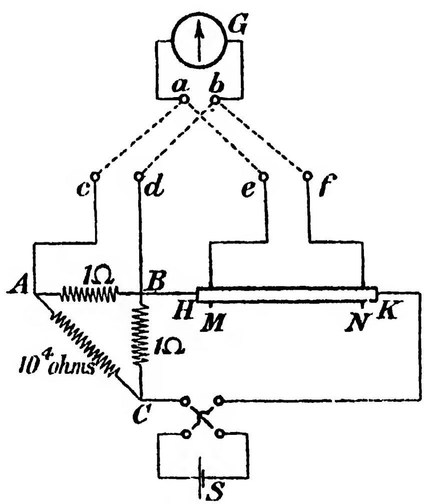


FIG. 46·9.—Determination of a Small Resistance.

$$\frac{Ir}{I \times 10^{-4}} = \frac{\theta_2}{\theta_1}, \text{ or } r = \frac{\theta_2}{\theta_1} . 10^{-4} \text{ ohm.}$$

Hence r may be found. Instead of using a galvanometer a potentiometer may be used to compare the p.ds. across AB and MN. If MN is of the order 10^{-2} ohm. a millivoltmeter may be used for this purpose, AC being replaced by a 100-ohm. coil. The current through AC and AB is then $\frac{1}{100}$nd part of that through MN.

The Variation of Resistance with Temperature.—By measuring the resistance of a wire at different temperatures it has been found that, in general, the resistance increases as the tempera-

ture is raised. For most pure metals the curve showing the variation of resistance with temperature measured on the gas scale is almost a perfect parabola, and for platinum CALLENDAR has shown that the variation of resistance with temperature is very accurately represented by the equation

$$R_\theta = R_0(1 + \alpha\theta + \beta\theta^2),$$

where θ is the temperature on a centigrade scale with air as the working substance, R_θ being the resistance at $\theta°$ C., and R_0 that at 0° C. To determine the constants α and β it is necessary to measure the resistance of a coil of platinum wire at three well-known temperatures, e.g. those of melting ice, steam at atmospheric pressure, and sulphur vapour. These are 0° C., 100° C. and 444·60° C. respectively, if in the last two instances the barometer reads 76 cm. of mercury under standard conditions. This last temperature has been determined accurately with the aid of a compensated gas thermometer [cf. p. 201].

The above equation may be written

$$\left(\frac{R_\theta - R_0}{R_0}\right)\frac{1}{\theta} = \alpha + \beta\theta.$$

Hence α and β may be calculated when the values of $\left(\frac{R_\theta - R_0}{R_0}\right)\frac{1}{\theta}$ of this equation are known at two different temperatures: 0° C. is excluded because the expression becomes indeterminate: of course a graphical method could be used by making observations on R_θ at several known temperatures.

On Centigrade Resistance Scales of Temperature.—If R_0 and R_{100} are the resistances of any given piece of wire at the temperatures of melting ice and of steam produced under standard conditions a centigrade resistance scale of temperature may be constructed by drawing two ordinates OA and BN, Fig. 46·10, to represent to scale R_0 and R_{100}, where the distance ON is 100 arbitrary units. To find the temperature ϕ, on this scale, corresponding to a resistance value R_ϕ, OH is constructed to represent R_ϕ, HK drawn parallel to ON to intersect AB in K, and KM drawn normal to ON to intersect this line in M. Then OM is ϕ units long, so that ϕ becomes known.

For temperatures measured on this scale of temperature, the relationship between R_ϕ and R_0 is a linear one and may be written

$$R_\phi = R_0(1 + \kappa\phi),$$

where κ is a constant, known as *the coefficient of increase of resistance with temperature as measured on the specified scale of temperature.* Now it so happens that for pure metals κ is very nearly equal to α in the expression $R_\theta = R_0(1 + \alpha\theta + \beta\theta^2)$, which means of course that

β is a small quantity: it is also found that $\kappa \doteqdot \alpha \doteqdot \frac{1}{273}$ deg.$^{-1}$ C. For alloys, κ and α, while still remaining nearly equal for any given alloy, vary considerably from one alloy to another and in the case of constantan and manganin tend to zero. It is for this reason that laboratory and standard resistance coils are constructed from these alloys.

To determine κ for iron, or nickel, a length of wire is wrapped on a wooden or mica frame and placed in a test-tube. Thick copper leads enable the coil to be connected to a P.O. box. The whole

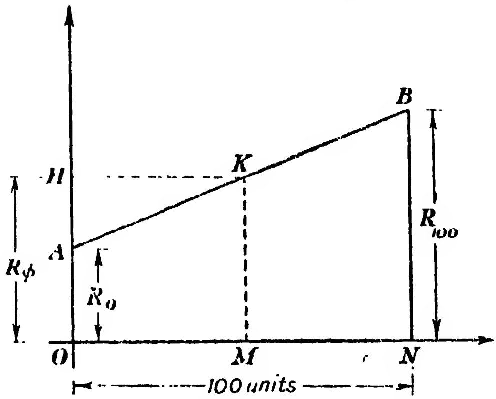


FIG. 46·10.—A Centigrade Resistance Scale of Temperature.

is placed in turn in melting ice and in steam and the resistance determined in each instance. The steam temperature is deduced from the barometric height and κ calculated from the equation

$$\kappa = \frac{R_{\phi_s} - R_0}{R_0\phi_s},$$

where ϕ_s is the steam temperature, calculated from the barometric height. No thermometer is used in this experiment, but a calibrated thermometer would be necessary if we wished to make a series of observations of R_θ at different temperatures θ and show that R_θ was very nearly a linear function of R_0 and θ.]

Platinum Resistance Thermometers.—The variation of resistance with temperature as a means of measuring temperature was first used by SIEMENS. His thermometer consisted of a platinum wire wound on a clay cylinder and mounted in an iron tube. The resistance of this thermometer in ice was not constant after it had been used at high temperatures, for the clay attacked the wire and gases passed through the iron causing the wire to become brittle. The physicists of his day therefore regarded the method as unpromising and it was not until about 1887, when CALLENDAR wound the wire on a mica frame without straining the wire, and

mounted the whole in a glass tube that this method of thermometry was developed. To-day it is one of the most reliable means of measuring temperatures from $-40°$ C. to $1{,}200°$ C.

A typical platinum resistance thermometer is indicated in Fig. 46·11 (*a*). The fine platinum wire is wound on a mica frame. This wire is joined by intermediate short lengths of thicker platinum wire to thick copper, silver or platinum leads. To compensate for the fact that the leads are at temperatures different from that of the

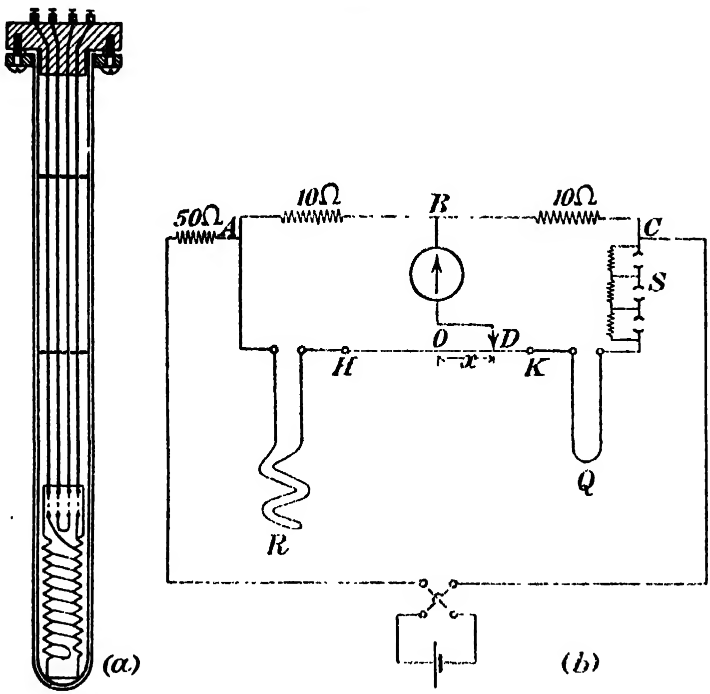


FIG. 46·11.—Platinum Thermometry (Pyrometry).

platinum spiral a pair of compensating leads is used. These are identical with the other leads and are joined to a short length of the same fine platinum wire through intermediate pieces of thicker platinum wire. The leads are held in position by an ebonite cap and mica washers and crosses; these also serve to insulate the leads from one another. The difference in resistance between each pair of leads and the wires connecting them depends only on the temperature of the platinum spiral; hence there is no troublesome and uncertain correction for stem exposure when such a thermometer is used. The above difference in resistance is measured on a CALLENDAR-GRIFFITHS bridge, the connections for which are shown

in Fig. 46·11 (*b*). Two 10-ohm coils are placed in the arms AB, BC. HK is a platinum wire corresponding to that in the Carey Foster bridge. The thermometer R is placed between A and H, while its compensating leads and a series of coils S of known resistance constitutes the fourth arm of the bridge. A 50-ohm resistance coil is placed in series with the battery to prevent a large current from being sent through the thermometer. When the bridge is balanced let us assume that D is x cm. from the *centre* O of the wire. Let ρ be the resistance per unit length of the wire. Then if $2l$ is the length of the bridge wire

$$R + \rho(l + x) = S + \rho(l - x)$$

since the ratio arms are equal. Hence $R = S - 2\rho x$.

The factor 2 in the above equation is troublesome when the bridge is used continually, so Callendar selected a bridge wire having a resistance 0·005 ohm. cm.$^{-1}$. Then $R = S - 0{\cdot}01 . x$.

The Platinum Resistance Scale of Temperature.—The centigrade resistance scale of temperature for platinum is very important, for it was used by Callendar to determine temperatures on the gas scale very accurately. Of course any metal may be used to define a centigrade resistance scale of temperature, but it is only for platinum that the relation between R_θ and θ is accurately parabolic and that α and β, in $R_\theta = R_0(1 + \alpha\theta + \beta\theta^2)$, do remain constant.

Since $R_\phi = R_0(1 + \kappa\phi)$, if R_0 and R_{100} are known,

$$\phi = \frac{R_\phi - R_0}{R_{100} - R_0} \times 100,$$

where ϕ is the temperature of the wire on the platinum resistance scale of temperature. $R_{100} - R_0$ is termed the fundamental interval of the thermometer and corresponds to 100° C.

To determine the temperature, θ, on the nitrogen centigrade gas scale corresponding to ϕ the constants α and β in the equation $R_\theta = R_0(1 + \alpha\theta + \beta\theta^2)$—this being the accurate relation between R_θ, R_0 and θ—are first determined as already explained.

Then

$$\phi = \left(\frac{\dfrac{R_\theta}{R_0} - 1}{\dfrac{R_{100}}{R_0} - 1}\right) \times 100 = \frac{\alpha\theta + \beta\theta^2}{100\alpha + 10{,}000\beta} \times 100$$

$$= \frac{\alpha\theta + \beta\theta^2}{\alpha + 100\beta}.$$

Hence

$$\theta - \phi = \frac{100\beta\theta - \beta\theta^2}{\alpha + 100\beta} = \frac{-\beta\theta(\theta - 100)}{\alpha + 100\beta} = d . \theta . (\theta - 100),$$

where d is termed the ***difference coefficient.*** Since β is negative, d is positive. It is equal to $1{\cdot}50 \times 10^{-4}$ for most samples of platinum.

Example.—$R_0 = 12{\cdot}784$ ohm., $R_{100} = 17{\cdot}765$ ohm., and $R_\theta = 25{\cdot}668$ ohm. Hence

$$\phi = \frac{25{\cdot}668 - 12{\cdot}784}{17{\cdot}765 - 12{\cdot}784} \times 100 = \frac{1288{\cdot}4}{4{\cdot}981} = 258{\cdot}7^\circ.$$

In using the difference formula we assume $\theta = \phi$ and obtain

$$\theta - \phi = 1{\cdot}5 \times 10^{-4} \times 258{\cdot}7 \times 158{\cdot}7 = 6{\cdot}2^\circ$$
$$\therefore\ \theta = 264{\cdot}9^\circ.$$

We now use this value in the difference equation and get

$$\theta - \phi = 1{\cdot}5 \times 10^{4} \times 264{\cdot}9 \times 164{\cdot}9 = 6{\cdot}5^\circ$$
$$\therefore\ \theta = 265{\cdot}2^\circ\ \mathrm{C}.$$

This process could be continued, but, in practice, it is seldom necessary to proceed beyond this stage.

The Comparison of Capacitances

Capacitance.—The practical unit of capacitance is the ***farad,*** and a condenser has a capacitance of one farad if a charge of one coulomb raises its potential by one volt.

For most purposes a farad is too large a unit; we therefore use the microfarad as a convenient unit of capacitance. It is denoted by the symbol μF and is equal to 1×10^{-6} farad.

A still smaller unit is the micro-microfarad ($\mu\mu$F); this is 1×10^{-12} farad. Campbell has recently suggested that 1$\mu\mu$F should be called a picofarad (1 pF). [It is ***approximately*** equal to the capacitance of a sphere whose radius is 1 cm., i.e. 1 pF $\doteqdot$ 1 e.s.u. of capacitance.]

Note on Electrical Units.—Let it be required to find:

(i) the amount of electricity in e.s.u. equivalent to 1 coulomb.
(ii) the capacitance in e.s.u. equivalent to 1 farad.

Each of these relations (and many others) can be obtained readily by considering expressions for energy; if the quantities concerned are expressed in (c.g.s.) electrostatic units, the energy will be measured in ergs; if expressed in practical units, the energy will be given in joules. Now 1 joule $\equiv 10^7$ erg., and we have to remember that 300 volt. $\equiv$ 1 e.s.u. of potential difference.

(i) The energy of a charged condenser is $\frac{1}{2}qv$. If the charge is 1 coulomb and the potential difference 1 volt, the energy stored is $\frac{1}{2} \times 1 \times 1 = 0{\cdot}5$ joule $= 0{\cdot}5 \times 10^7$ erg.

Let 1 coulomb $= \alpha$ e.s.u. of charge. Then the energy of the above condenser is $(\frac{1}{2}.\alpha \times \frac{1}{300})$ erg. Hence

$$\frac{\alpha}{300} = 10^7 \text{ or } \alpha = 3 \times 10^9,$$

i.e. 1 coulomb $\equiv 3 \times 10^9$ (c.g.s.) e.s.u. of charge.

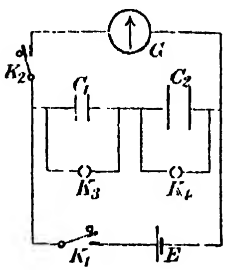
From this it follows that

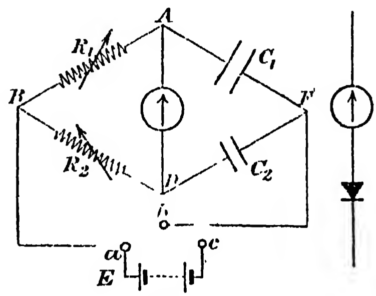
$$1 \text{ e.m.u. of quantity} \equiv 3 \times 10^{10} \text{ e.s.u. of charge.}$$

[NOTE.—3×10^{10} cm.sec.$^{-1}$ is the velocity of light in a vacuum and the last relation we have obtained is a consequence of the electromagnetic theory of light.]

(ii) $$1 \text{ farad} = \frac{1 \text{ coulomb}}{1 \text{ volt}} = \frac{3 \times 10^9 \text{ e.s.u. of charge}}{(\frac{1}{300}) \text{ e.s.u. of potential}}$$
$$= 9 \times 10^{11} \text{ e.s.u. of capacitance.}$$

The above result may be obtained directly from energy considerations as follows. The energy of a charged condenser is $\frac{1}{2}CV^2$. If $C = 1$ farad, and $V = 1$ volt, the energy is 0·5 joule.





(a) (b)

FIG. 46·12. FIG. 46·13.

Comparison of Capacities.

Let 1 farad $= \beta$ e.s.u. of capacitance. Then the energy of the above condenser is $\frac{1}{2}\beta \times (\frac{1}{300})^2$ erg.

$$\therefore \beta \times (\tfrac{1}{300})^2 = 10^7, \qquad \text{or } \beta = 9 \times 10^{11},$$

i.e. 1 farad $= 9 \times 10^{11}$ e.s.u. of capacitance $= 9 \times 10^{11}$ cm.

$$1\mu F \equiv 9 \times 10^5 \text{ e.s.u.}$$
$$1\mu\mu F \equiv 0{\cdot}9 \text{ e.s.u.}$$

The Comparison of Condensers.—(*a*) The circuit necessary for this is indicated in Fig. 46·12. To commence the experiment all the keys are open. To charge C_1 the key K_4 is closed, i.e. the condenser C_2 is short circuited, and then K_1 is closed. By opening K_1 and closing K_2 the condenser is discharged through the galvanometer. Let σ_1 be its throw. If Q_1 is the quantity of electricity which has passed, and E the e.m.f. of the battery, $Q_1 = C_1E$. K_3 is then closed and the experiment repeated with C_2. Then $Q_2 = C_2E$. But $\frac{Q_1}{Q_2} = \frac{\sigma_1}{\sigma_2}$, so that $\frac{C_1}{C_2} = \frac{\sigma_1}{\sigma_2}$.

(*b*) **Comparison of the Capacitances of Condensers by de Sauty's Method.**—The two condensers to be compared, C_1 and C_2,

and two adjustable resistances, R_1 and R_2, are arranged to form the four arms of a network similar to the Wheatstone bridge arrangement of resistances—cf. Fig. 46·13 (*a*). A high-resistance galvanometer is connected across the points AD; E is a battery of about 10 volts. This is connected to the bridge and mercury cups *a*, *c*, as indicated. When a connecting wire across *ab* is removed and placed across *bc* the condensers are charged, the potential difference across each condenser being equal to V, the e.m.f. of the battery. The resistances R_1 and R_2, which should be large, are adjusted until there is no kick of the galvanometer when the condensers are charged or discharged. When this condition is satisfied,

$$\frac{C_1}{C_2} = \frac{R_2}{R_1}.$$

Proof.—Let Q_1 and Q_2 be the charges on the (positive) plates of the condensers C_1 and C_2 respectively. Then

$$Q_1 = C_1V \text{ and } Q_2 = C_2V.$$

To obtain the fraction of the charge on C_1 passing through G on discharge, let I_1 be the instantaneous value of the current in the wire connecting A to the positive plate of C_1. To reach B this may travel either via R_1, or via G and R_2. It will divide itself so that the instantaneous potential difference across R_1 is equal to the sum of the instantaneous potential differences across G and R_2. The instantaneous current through G is therefore $\frac{R_1}{R_1 + R_2 + G} \cdot I_1$, where G is the resistance of the galvanometer. If this current lasts for time δt, the quantity of electricity passing is $\frac{R_1}{R_1 + R_2 + G} \cdot I_1 \cdot \delta t$. If τ is the time required for the discharge, the total charge from C_1 passing through G is

$$\frac{R_1}{R_1 + R_2 + G} \cdot \int_0^\tau I_1 dt = \frac{R_1}{R_1 + R_2 + G} \cdot Q_1$$

$$= \frac{R_1}{R_1 + R_2 + G} \cdot C_1V.$$

Similarly, the fraction of the charge on C_2 which passes through G on discharge is

$$\frac{R_2}{R_1 + R_2 + G} \cdot C_2V.$$

Since these charges pass in opposite directions through the galvanometer, and the total quantity passing is zero when the bridge is balanced, we have

$$R_1C_1 = R_2C_2.$$

Hence

$$\frac{C_1}{C_2} = \frac{R_2}{R_1}.$$

Instead of using a battery and a tapping key, alternating current may be employed. A crystal, such as those used in some wireless receiving sets, must then be placed in series with the galvanometer, or the galvanometer may be replaced by a pair of phones. It is very essential in these experiments that the resistance coils should be non-

inductive [cf. p. 962]. [The method of indicating that a crystal is in series with the galvanometer is shown in Fig. 46·13 (*b*).]

The Comparison of e.m.fs.—To use a ballistic galvanometer for this purpose a condenser is charged from one cell and then discharged. The same condenser is connected in turn to the other cells and the throws of the galvanometer are observed. Since $Q = CV = \kappa\sigma$, where σ is the throw, and κ a constant, it follows that the e.m.fs. are directly proportional to the throws.

The Internal Resistance of a Voltaic Cell—Lodge's Method.—The cell is first connected to a condenser and the throw, σ_1, of a galvanometer observed when the condenser is discharged through it. A known resistance is then placed across the cell so that it is supplying a current. If the condenser is connected to the cell and the throw, σ_2, due to its discharge observed, it will be found to be smaller than the first throw. This is because the condenser has only acquired a p.d. equal to the e.m.f. of the cell on closed circuit under the above conditions. We have already seen

$$\frac{E - V_c}{B} = \frac{V_c}{R}.$$

Since E and V_c are proportional to σ_1 and σ_2 respectively, we have

$$\frac{\sigma_1 - \sigma_2}{B} = \frac{\sigma_2}{R}.$$

Hence B may be determined.

The Grouping of Cells.—Let us consider three different methods of arranging N cells, each of e.m.f. E and internal resistance r to send a current through a resistance R.

(i) ***The Cells in Series.*** In this instance the total e.m.f. in the circuit is NE: the current is

$$I = \frac{NE}{Nr + R}.$$

(ii) ***The Cells in Parallel.*** The e.m.f. of the battery is identical with that of one cell, since the e.m.f. of a cell is independent of the size of the plates. The internal resistance, B, of the battery is the resistance equivalent to N resistors, each of resistance r, in parallel, viz.

$$\frac{1}{B} = \frac{1}{r} + \frac{1}{r} + \frac{1}{r} + \ldots \text{ to N terms} = \frac{N}{r}$$

$$\text{or } B = \frac{r}{N}.$$

$$\therefore I = \frac{E}{R + \frac{r}{N}}.$$

(iii) ***The Cells in Series-Parallel.*** Suppose that the N cells are arranged in n rows so that each row contains $\frac{N}{n}$ cells, cf. Fig. 46·14 (a). The internal resistance of the cells in each row is $\frac{N}{n}.r$. Since the rows are in parallel the internal resistance, B, of the battery is given by

$$\frac{1}{B}=\frac{n}{N}\cdot\frac{1}{r}+\frac{n}{N}\cdot\frac{1}{r}+\ldots \text{ to } n \text{ terms}=\frac{n^2}{N}\frac{1}{r}.$$

$$\therefore B=\frac{Nr}{n^2}.$$

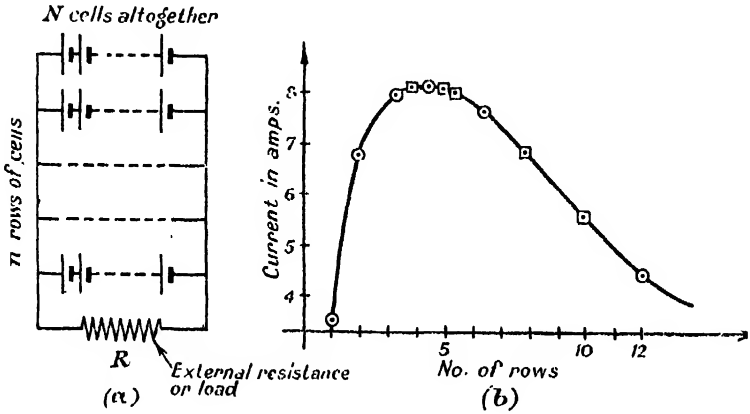


FIG. 46·14.

The e.m.f. of the battery is equal to that of $\frac{N}{n}$ cells, viz. $\frac{NE}{n}$.

$$\therefore I=\frac{\frac{NE}{n}}{\frac{Nr}{n^2}+R}=\frac{NE}{\frac{Nr}{n}+nR}.$$

Now the current is a maximum when the denominator in the above fraction is a minimum. If $Z=\frac{Nr}{n}+nR$, we have $\frac{dZ}{dn}=0$, for a maximum or a minimum. Now

$$\frac{dZ}{dn}=-\frac{Nr}{n^2}+R=0,$$

$$\text{or } n=\sqrt{\frac{Nr}{R}},$$

since $\frac{dZ}{dn}=0$ for a maximum or a minimum.

Physical considerations, or the fact that $\dfrac{d^2Z}{dn^2}$ is negative when $n = \sqrt{\dfrac{Nr}{R}}$, show that the current is a maximum when $n = \sqrt{\dfrac{Nr}{R}}$.

Example.—Twelve cells each of e.m.f. 2 volts and of internal resistance 0·5 ohm are available. Determine, graphically, the largest current which may be sent through a resistor of 0·375 ohm resistance.

The possible values for n, the number of cells in a row, are 1, 2, 3, 4, 6, 12. The general expression for the current is $I = \dfrac{24}{\dfrac{12 \times 0\cdot5}{n} + n(\frac{3}{8})} = \dfrac{64n}{16 + n^2}$, so that the currents corresponding to the above possible values for n are 3·77, 6·40, 7·84, 8·00, 7·39, 4·8 amperes. The graph, shown in Fig. 46·14 (b), where the above points are indicated by $\odot$'s, suggests that I is a maximum when $n = 4$. To get the correct shape of the graph we can put $n = 3\cdot5$, $4\cdot5$, 5, 8, 10, etc. [shown $\boxdot$], for although such numbers are impossible physically, they may be used to plot the graph.

The maximum current is 8 amp.

EXAMPLES XLVI

1.—You are supplied with an ammeter and voltmeter. How would you proceed to measure the resistance of an electric lamp? A potential difference of 31 volts is sufficient to cause a current of 0·127 ampere to pass through it. When the potential is raised to 109 volts, the current is 0·225 amp. Compare the resistance of the lamp when the p.d. across it is 109 volts with its resistance in the first instance.

2.—A potentiometer wire is 100 cm. long. A constant p.d. is maintained across it. Two cells, A and B, are connected in series first to support one another and then in opposition. The balance points are 60·2 cm. and 12·3 cm. from the same end of the wire when the two arrangements are compared. Calculate the ratio of the e.m.fs. of the cells.

3.—The potential difference across a 2 ohm. coil is measured by means of a potentiometer. The balance point is at 57·8 cm. When a Daniell cell is used the balance point is at 30·7 cm. What is the current in the 2 ohm. coil? [e.m.f. of cell = 1·08 volt.]

4.—How is the electrical resistance of a circuit defined? How would you measure the resistance of (a) an accumulator cell, (b) a galvanometer?

5.—Describe and explain how you would test the resistance of a 0·1 ohm. shunt (a) using a standard 0·1 ohm. coil, (b) when the smallest available standard coil is 1 ohm.

6.—Lengths of wire having resistances 1 ohm., 1 ohm., and 100 ohms. form the sides AB, BC, and CA of a triangle ABC. One end of a thick copper wire X is connected to B and its other end to a variable resistance and one pole of a battery. The second pole of this is connected to C. When a high-resistance galvanometer is connected in turn across AB and across X deflexions of 33·5 and 32·0 divisions are obtained. If the length of the wire is 111·5 cm. and its diameter is 1·64 mm., calculate the resistivity of copper.

7.—Explain how, with the aid of a tangent galvanometer, a standard 1 ohm. coil, a potientiometer, adjustable resistances, cells, and other apparatus usually found in a laboratory, you would determine the relation between the electromagnetic unit of current and the ampere. [No ammeter or voltmeter is allowed.]

8.—Describe the essential features, and the principle of the action of a quadrant electrometer.

Explain how you would use this instrument to compare the resistances of two metal wires. What conditions limit the magnitude of the resistances which can be compared in this way ?

9.*—Write a short essay on the measurement of temperature above the range of the mercury thermometer. (L. '31.)

10.—Two condensers of capacitances 0·1 μF. and 0·02 μF. are connected in series across a battery of e.m.f. 24 V. They are then insulated and connected in parallel. What is the loss of energy in ergs ?

11.—Describe a Post Office box and explain how you would use it to measure the coefficient of increase of resistance with temperature for nickel.

In a metre-bridge arrangement for the comparison of resistances, the two coils are made of copper and each has a resistance of 1 ohm. at 0° C. If one coil is maintained at 0° C. and the other at 50° C., what resistance must be put as a shunt across the hotter coil in order that the balance of the bridge may not be disturbed ? Calculate, also, the displacement of the slide contact necessary to rebalance the bridge without using the shunt. The temperature coefficient of resistance for copper may be taken as 0·004 deg.$^{-1}$ C.

12.—A condenser of capacitance 0·5 μF. is charged to a p.d. of 10 volt. and then discharged through a galvanometer. This process is repeated 100 times a second and in such a way that only the discharge current passes through the galvanometer. Draw a circuit diagram of the way in which this may be accomplished, and calculate the deflexion of the galvanometer whose resistance is 16 ohm. and current sensitivity 0·1 division per micro-ampere.

If, instead of the above arrangement, the same galvanometer were placed in series with the 10 volt. battery and a resistance R ohm., so that the same deflexion as before was obtained, calculate the value of R. Why may the resistance of the galvanometer be neglected in each part of the above experiment ?

[1 μF. = 10^{-6} farad.]

*This question should only be attempted after reading the chapters on thermoelectricity, etc.

CHAPTER XLVII

THE PHENOMENA OF ELECTROLYSIS

Electronic and Electrolytic Conductors.—Conductors of electricity may be divided into two classes—***metallic*** or ***electronic*** conductors such as copper, brass, graphite, and certain oxides, and ***electrolytic*** conductors or ***electrolytes*** such as aqueous solutions of inorganic acids and salts, and fused salts. In both instances the passage of electricity is accompanied by magnetic and heating effects, but there is an essential difference between them also. This difference is to be found in the fact that when electricity is passed through an electrolyte it is always associated with a transference of matter. This transportation only occurs in the electrolyte, for when the current leaves it the matter can no longer accompany it and must therefore be liberated in the electrolyte. In metallic conduction we do not observe any chemical change in the conductor as the result of the passage of electricity through it —the current is not associated with any motion of the substance of the conductor. The current in this instance is carried by electrons—negatively charged particles of very minute mass.

Electrolysis.—If an electric lamp and two copper plates dipping into distilled water are connected to the mains the lamp does not glow. But, if one or two drops of concentrated sulphuric acid are added to the water, the lamp immediately glows. In addition, gases will be observed to be liberated at the copper plates. If the experiment is repeated using a strong solution of copper sulphate instead of the acid it will be found that gas is liberated at one plate whereas copper is deposited at the other.

Electro-positive and Electro-negative Elements.—Whenever a compound, consisting of two chemical elements, is resolved by the passage of an electric current, the elements are liberated at the electrodes: whether these elements actually appear at the metal plates or not is determined by their chemical affinity for the solute or for the electrodes. The elements liberated at the ***negative electrode*** or ***cathode*** are said to be ***electro-positive***; those set free at the ***positive electrode*** or ***anode*** constitute the ***electro-negative elements.*** The former class comprises all the metals and hydrogen; the non-metallic elements are generally electro-negative. Chemically there is a stronger tendency for electro-

positive elements to combine with electro-negative elements than there is for either class to form compounds among themselves.

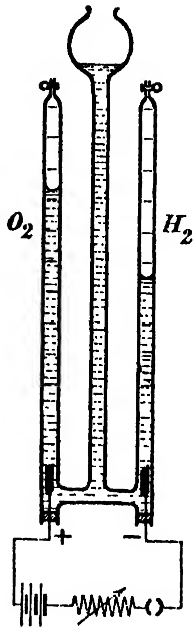


FIG. 47·1.—Electrolysis of Acidulated Water.

The Electrolysis of Dilute Sulphuric Acid [Acidulated water].—For this experiment the apparatus shown in Fig. 47·1 is required. It consists of three glass tubes joined together at their lower extremities; two of these tubes are furnished with stop-cocks, whilst the central one is widened at its upper end. The two rectangular electrodes are made of platinum and are connected to a suitable battery. The two outer limbs are completely filled with the acid, whilst the third limb and its reservoir serve to hold the liquid which is displaced from the other limbs when the current flows through the electrolyte. This displacement is caused by the gases which are liberated at the electrodes. Graduations on each limb enable the volume of gas to be ascertained and, in this experiment, independently of the magnitude of the current or the time for which it has passed, it will be found that the ratio of the volumes of the gases in the two limbs is 2 : 1. The application of a lighted match, and the subsequent blue flame, indicates that the larger volume of gas is hydrogen; the other gas is oxygen [test with a glowing piece of wood].

If an aqueous solution of baryta, $Ba(OH)_2$, is employed the products of the electrolysis are still hydrogen and oxygen, the ratio of their volumes being as before. This particular solution is mentioned because the products of the electrolysis are exceptionally pure, so that pure samples of hydrogen and oxygen are easily obtained if pure barium hydroxide and distilled water are used in the voltameter.

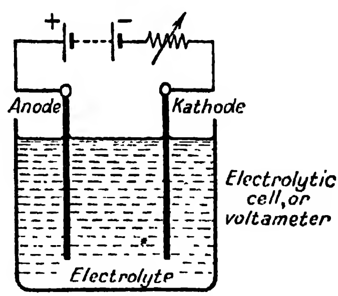


FIG. 47·2.

So far we have only obtained information concerning the products of electrolysis. For the purpose of ascertaining any changes which may have occurred in the electrolyte itself the apparatus shown in Fig. 47·2 may be used. After the current has passed for some time samples of

the electrolyte from regions near the anode and cathode are withdrawn and the amount of acid in them estimated by chemical methods. It will be found that the concentration has increased in the neighbourhood of the anode [where oxygen is evolved during the electrolysis] and decreased round the cathode. If, however, the experiment is repeated and the total amount of acid in the electrolyte determined it will be found to be constant. The above effects (which are typical of all electrolytes) may be summarized thus :—

(*a*) Water has been decomposed into its elements which appear separated from one another.

(*b*) Although the total amount of acid has remained constant, local changes in concentration have occurred.

The Electrolysis of Copper Sulphate Solution.—(*a*) ***Soluble Electrodes.*** If two copper plates are immersed in an aqueous solution of copper sulphate and connected to a battery, copper will be deposited on the cathode. At the same time the copper anode gradually dissolves, forming copper sulphate which replenishes the solution so that the total amount of copper sulphate in solution is constant.

(*b*) ***Insoluble Electrodes.*** If, in the above experiment, the copper electrodes are replaced by some of platinum, copper will still be deposited at the cathode but bubbles of oxygen will be evolved at the anode and the copper sulphate in solution gradually diminishes. In its place there will appear sulphuric acid. If the experiment is continued until all the copper sulphate has been removed electrolysis will not cease, for the electrolyte will now be acidulated water and the products hydrogen and oxygen. [This continuation of the electrolysis will only proceed if the p.d. across the electrolyte is greater than 1·67 volts—solely to overcome the polarization or back e.m.f. in the electrolytic cell.]

Faraday's Laws of Electrolysis.—Let us suppose that three vessels, A, B and C, Fig. 47·3, contain copper sulphate solution. In each one a sheet of copper forms the cathode, whilst the anode is of platinum and dips underneath an inverted test-tube which is filled with water. A and B are connected together in parallel, this combination then being placed in series with C, a battery, and an ammeter. The current is allowed to flow, and it will be observed that the total quantity of oxygen collected in the test-tubes in A and B is equal to the amount in C. Similarly, the total mass of copper deposited on the cathodes in A and B is equal to the mass deposited on the third cathode. Furthermore, if the experiment is continued for different lengths of time, it will always be found that the quantity of copper or oxygen liberated is directly proportional to the time. These and similar results are summarized

in Faraday's first law of electrolysis which states : ***The mass of any individual product liberated in electrolysis is directly proportional to the quantity of electricitty which has passed through the electrolyte.***

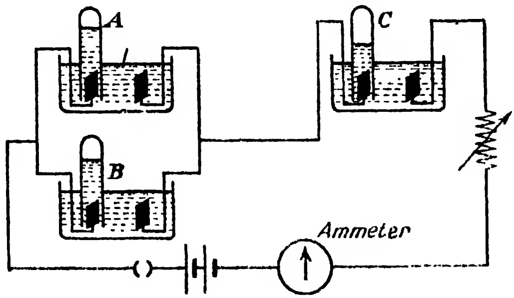


Fig. 47·3.

If aqueous solutions of copper sulphate, silver nitrate, and sulphuric acid are connected in series and the same current is passed through each cell, experiment shows that the mass of copper deposited is to the mass of silver as 31·8 : 108, i.e. for every 31·8 gm. of copper deposited there are 108 gm. of silver set free. Under the same circumstances 1·008 gm. of hydrogen will have been collected. Now these numbers are the chemical ***equivalents*** of the respective elements, so that these and similar results may be summarized thus: ***The masses of the different products liberated in electrolysis by the passage of equal quantities of electricity are proportional to their chemical equivalents, where the chemical equivalent of an element is defined as the mass of that element which will combine with, or displace, 8 gm. of oxygen in a chemical reaction.*** This is a statement of Faraday's second law of electrolysis.

Electrochemical Equivalents.—Since when a steady current of I amperes flows for t seconds, the quantity of electricity passing is It, we may express the first law of electrolysis by the equation

$$m = z\mathrm{I}t,$$

where m is the mass of substance liberated at an electrode and z is a constant. If m is expressed in grams, and the substance liberated is an element, z is known as the ***electrochemical equivalent*** of that element. Thus the statement that the electrochemical equivalent of silver is 0·001118 gm.coulomb.$^{-1}$ signifies that this mass of silver is deposited by one ***coulomb,*** i.e. by a steady current of 1 ampere flowing for 1 second. A similar meaning is given to the statement that the electrochemical equivalent of copper is 0·000329 gm.coulomb.$^{-1}$, and such facts have been

established so well that they provide a ready and accurate means of determining the corrections which may have to be applied to the readings of a given ammeter or voltmeter. The latter instrument must be shunted across a suitable coil of known resistance [cf. p. 854].

The quantity of electricity necessary to liberate one gram-equivalent of a substance by electrolysis is termed a ***faraday.*** Accurate experiments have shown that

$$1 \text{ faraday} = 96{,}490 \text{ coulombs} = 26{\cdot}8 \text{ ampere-hours.}$$

Voltameters.—To determine the electrochemical equivalent of a substance or, when this is known, to measure a quantity of electricity, voltameters are used. A silver voltameter due to RICHARDS is indicated in Fig. 47·4. The cathode is a platinum crucible containing a freshly prepared 10 per cent. $AgNO_3$ solution. The anode consists of a pure silver rod wrapped in filter paper and immersed in $AgNO_3$ solution. The anode and cathode compartments are separated by a porous pot to diminish certain small disturbing factors. Richards found this to be necessary in accurate work.

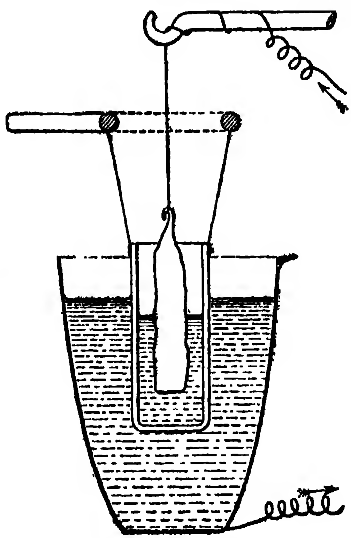


FIG. 47·4.—Silver Voltameter.

For students' use and when results of the greatest accuracy are not required a copper voltameter or coulometer is employed. Three sheets of copper are placed in the electrolyte, the two outer ones constituting the anode while the inner one is the cathode. If ordinary sheet copper is used the electrodes should be enclosed in paper bags to prevent impurities from straying into the electrolyte. The electrolyte may be stirred by passing a stream of hydrogen through it. This latter precaution is only necessary when the voltameter is in use for considerable periods.

Electricity Meters.—Large quantities of electricity may be measured by a meter due to WRIGHT—cf. Fig. 47·5. Only a small known fraction of the current passes through the meter. The anode is mercury contained at a constant level in A by means of a reservoir B. The mercury in A is in the form of a ring. A piece of iridium foil C forms the cathode. The electrolyte is a solution of mercuric iodide in aqueous potassium iodide. At the cathode the product of electro-

lysis is mercury, and since this does not amalgamate with iridium, drops of mercury fall through the funnel into the U-tube D. This tube is graduated so that the volume of mercury may be read off. This volume is a measure of the quantity of electricity which has passed in the main circuit. When D is full the mercury siphons over and is collected at E: another scale gives the volume of mercury collecting there. The amount of mercury in solution remains constant, for the anodic mercury dissolves. When the instrument is inverted the mercury flows from E to B and the instrument is again ready for use. F is a fence of glass to prevent mercury passing from the anode to E should it receive an accidental mechanical shock.

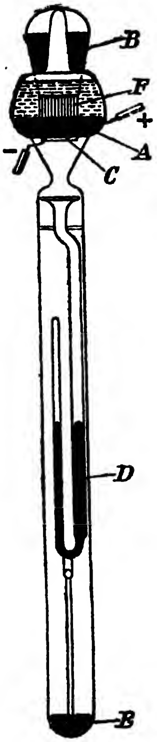


FIG. 47·5. — Wright's Electricity Meter.

Ohm's Law for Electrolytic Conduction.—To obtain the characteristic of an electrolyte the apparatus shown in Fig. 47·6 may be used. The electrolyte [acidulated water] is contained in a cell of the type indicated, the connecting tube having a diameter of 1 cm. and the electrodes being platinum coated with platinum black. The current through the cell is changed by altering the sliding resistance, A, which is used as a potential divider, the current being measured by a high-resistance millivoltmeter, MV, shunted across a 10-ohm coil. The potential difference across the electrolytic cell is measured on a potentiometer. Since the resistance of the wire in this is small compared with that of the cell, it is necessary to use a high resistance galvanometer, G, or else place 5,000 ohms in series with it, for otherwise the current in the main circuit is considerably disturbed while the point of balance on the potentiometer is being found. Corresponding readings of the current through the cell and the p.d. across it are observed. The curve shown in Fig. 47·7 represents the results thus obtained.

This characteristic is a straight line, but it does not pass through the origin of co-ordinates. This is because the back e.m.f., v, in the cell must be overcome before any current will pass through it. If V is the applied p.d. as measured by the potentiometer, then the above graph shows $\left(\frac{V-v}{I}\right)$ to be constant. This does not mean that Ohm's law is not true, for the expression merely takes this particular form since the potentiometer measures the p.d. necessary to overcome the back e.m.f. together with that necessary to send the current through the cell. The resistance of

the liquid column is $\left(\frac{V - v}{I}\right)$. We can only speak of Ohm's law in reference to electrolytes, i.e. there will only be a linear relationship between the available potential difference $(V - v)$ and the current

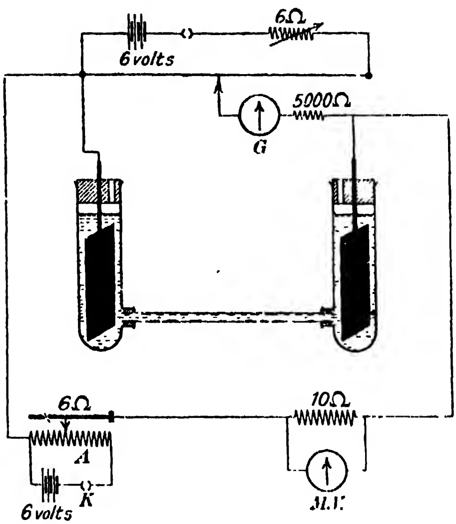


FIG. 47·6.—Verification of Ohm's Law for an Electrolyte (Acidulated Water).

I, if the current passed through the electrolyte is so small and exists only for such a short time that the concentration of the electrolyte is not altered, for otherwise we should violate one essen-

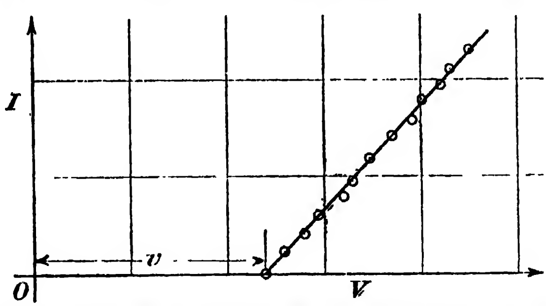


FIG. 47·7.—Characteristic Curve for an Electrolyte (Acidulated Water).

tial condition of Ohm's law, viz., the state of the conductor must remain constant.

With copper electrodes and an aqueous solution of copper sulphate the characteristic obtained in the above manner is a straight

line through the origin because the back e.m.f. in this instance is zero.

For success in these experiments it is advisable to begin with the largest value of the current and gradually reduce it to zero. If this is done the back e.m.f. in the cell soon attains its maximum value and conditions become steady.

The Resistance of Electrolytes.—A direct current cannot be employed in general for the determination of the resistivity of an electrolyte because of the polarization occurring at the electrodes. KOHLRAUSCH used alternating current and designed the following

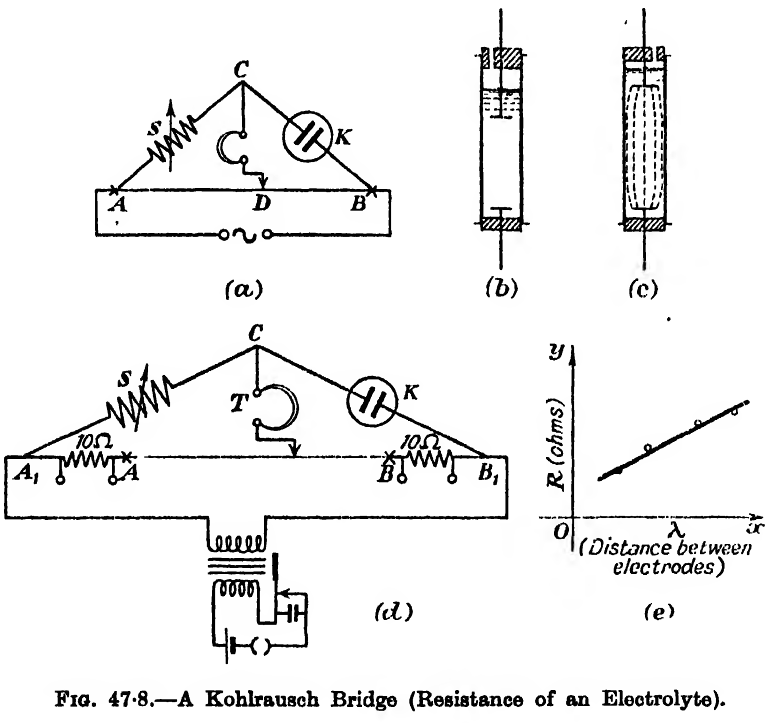


FIG. 47·8.—A Kohlrausch Bridge (Resistance of an Electrolyte).

bridge which has been named after him. Fig. 47·8 (*a*) is a diagram of the circuit. AB is a uniform wire, S an adjustable known resistance, and K the cell containing electrolyte. Intermittent or alternating current is supplied as indicated. Now Kohlrausch showed that the frequency of the supply must be large—50 cycle.sec.$^{-1}$ or more—if the above effects are to be eliminated. A small induction coil is therefore often used. The electrolyte is contained in a glass tube provided with rubber bungs through which the electrodes pass—cf. Fig. 47·8 (*b*). Reference to the bridge shown in Fig.

47·8 (*a*) shows that it is similar to a Wheatstone bridge : only since alternating current is used a telephone replaces the usual galvanometer. As the jockey moves along the bridge wire the sound in the 'phones varies. When this sound is a minimum—complete silence is seldom obtained—we have the ordinary Wheatstone bridge relationship, viz.,

$$\frac{R}{S} = \frac{l_1}{l_2},$$

where R is the resistance of the electrolyte, S a known resistance, and l_1 and l_2 the lengths of the corresponding parts of the bridge wire. [If the sound is a minimum over a range of the bridge wire instead of at some definite point, the extent of this range should be recorded and its centre considered to be the true balance point.]

In practice the above simple circuit is not sensitive since the wire AB almost short-circuits the secondary of the induction coil—under such circumstances the rate of supply of energy to the bridge is small. To increase this AB may be made from very fine wire, but this is not desirable for such wire is easily damaged. The difficulty is overcome by increasing the effective length of AB by placing known resistance coils at its ends as shown in Fig. 47·8 (*d*). Generally they are equal—say 10 Ω. To begin the experiment these coils are short-circuited and the jockey placed near to the middle of AB. S is adjusted until the bridge is balanced. The short-circuiting pieces of wire are then removed and the balance point accurately determined. This procedure is necessary to ensure that the balance point shall be on AB. Then

$$\frac{S}{R} = \frac{10 + \text{resistance of AD}}{10 + \text{resistance of DB}},$$

so that it is necessary to know the resistance per unit length of the wire AB.

Now, in general, $R = \chi \cdot \frac{l}{A}$, where the symbols have their usual meanings. In this instance χ cannot be deduced directly because the electrodes do not fit the tube and the lines of current flow are not linear—cf. Fig. 47·8 (*c*). To obtain a value for χ, R is determined for various distances, λ, between the electrodes. The relationship is a linear one if λ is not less than about 1·5 diameters of the tube, i.e. $R = \frac{\chi \cdot \lambda}{A} + c$, where c is a constant—cf. Fig. 47·8 (*e*). The slope of the line gives $\frac{\chi}{A}$, so that χ may be obtained.

When this bridge is used to determine the resistivity of a copper

sulphate solution, copper electrodes may be employed. [In all other cases platinum electrodes should be used.] They must be cleaned with nitric acid and thoroughly washed with distilled water before use. When the first accurate balance point has been obtained the electrodes should be treated in the above manner and then the final observations made. In this way the sharpness of the position for minimum sound in the 'phones is greatly enhanced.

An alternative cell is shown in Fig. 47·6. A glass tube of uniform width is held by rubber bungs in the manner indicated. The electrodes are fixed in position relatively to the supporting tubes. Now the resistance actually measured when such a cell is filled with electrolyte is the resistance offered by the electrolyte in the glass tube together with a resistance arising from the presence of the liquid between the electrodes and the ends of the uniform glass tube. This latter resistance may be eliminated by repeating the experiment with a portion of the tube removed, the distances from the ends of the tube to the electrodes being constant. The difference between the two resistances thus determined gives the resistance of the electrolyte in the portion of the tube cut off. If the mean diameter and length of this portion are known the resistivity of the solution may be calculated.

Electrolytic Dissociation Theory.—The two salient facts which any satisfactory theory of electrolysis must explain are (*a*) the transfer of electricity and (*b*) the transport of matter through the electrolyte. Following on the pioneer work of Grothuss and Clausius, in 1887, Arrhenius brought forward his theory of electrolytic dissociation. According to this, the conductivity of an electrolyte is attributable to the presence of free ions in the solution. These appear spontaneously whenever solution takes place. In electrolysis they consist of atoms, or groups of atoms, carrying one or more elementary electric charges. If the charge is positive the ion is called a ***cation*** since it moves towards the cathode (— *ve*) under the influence of an electromotive force. Similarly the ***anion*** is the ion with a negative charge.[1] It is necessary to assume that these ions do appear immediately solution occurs since Ohm's law is true for electrolytes. If Ohm's law were not valid, it would imply that work was necessary to split up the molecules of the dissolved substance before conduction in such an instance could occur. Thus the p.d. across an electrolyte would be greater than that implied by Ohm's law. Moreover, the theory assumes that there is a constant interchange of ions between the molecules, and at any instant there is a large number of ions in the act of migrating

[1] This charge is equal to that on an electron—the so-called elementary charge. The charge on the positive ion of a monovalent element has a charge numerically equal to that of an electron.

from one molecule to another. Thus in an aqueous solution of copper sulphate, some of the sulphate molecules are broken up into Cu ions (cations) and SO_4 ions (anions). The copper ions carry two positive charges, a fact which is symbolized thus : $Cu^{\cdot\cdot}$. The sulphate ion is denoted by SO_4''. These ions are free to move at random throughout the cell before the e.m.f. is applied. When it is applied a directive force is exerted upon them, the anions (—) moving towards the anode (+) and the cations (+) to the cathode (—). Thus, when a solution of copper sulphate is electrolysed the copper cations move to the cathode where they lose their charges and copper is deposited. On the other hand, the sulphions (—) move to the anode and lose their charges. If this is made of copper the free SO_4 radical combines with the copper electrode and copper sulphate passes into solution. If, however, platinum electrodes are used copper is deposited on the cathode, but the free SO_4 radicals are unable to combine with platinum so that they react with the water as follows :—

$$2SO_4'' + 2H_2O \rightarrow 2H_2SO_4 + O_2\uparrow + \text{four negative charges which pass from the electrolyte.}$$

The electrolytic decomposition of dilute sulphuric acid is explained by supposing that the free ions are $H^{\cdot}$ and SO_4''. The cations (+) are directed under the influence of the electric field towards the cathode (—) where their charges are lost and they combine to give bubbles of hydrogen. The sulphions are directed in the opposite direction and, with platinum electrodes, eventually give rise to the formation of oxygen as previously explained. The net result is the decomposition of water.

$$H_2SO_4 \rightarrow 2H^{\cdot} + SO_4''$$
$$2H^{\cdot} + 2\,\varepsilon \rightarrow 2H \rightarrow H_2\uparrow$$
$$2[SO_4'' + H_2O] + 4\,\varepsilon \rightarrow 2H_2SO_4 + O_2\uparrow$$

Experiment.—The passage of sodium ions through glass may be shown in a very striking manner. L, Fig. 47·9, is a carbon filament lamp whose filament is raised to incandescence by means of a battery AB. This lamp is partly immersed in an iron trough containing molten sodium nitrate. The positive pole of the battery AB is connected to the negative pole of a high-tension battery CD, the positive pole of which is joined to the trough. A galvanometer, G, shunted by a resistance, S, if necessary, indicates the passage of a current in the circuit LGDCL. After about one hour the walls of the lamp are covered on the inside with a deposit of metallic sodium. To explain this we have only to consider the positive sodium ions which move under the influence of the electric field between the trough and the filament towards the latter. They pass through the glass [which contains sodium] into the bulb. Here they come into contact with electrons which are continually shot off from the hot filament so that their positive charges are lost and they become normal sodium atoms. These move to the cooler parts of the bulb where they are deposited.

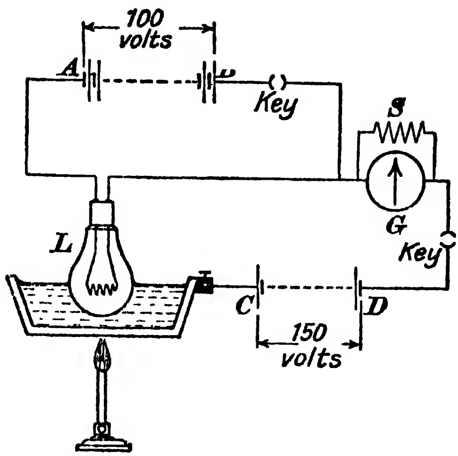


FIG. 47·9.—Passage of Sodium Ions through Glass.

Definitions.—The ***gram-equivalent*** of a substance is a quantity in grammes equal to its chemical equivalent. Thus, the atomic weight of silver is 108—its valency being unity. Its chemical equivalent is therefore 108. A gram-equivalent of silver is therefore 108 gm. Copper is bivalent and has an atomic weight 63: its chemical equivalent is 31·5 and a gram-equivalent of copper is therefore 31·5 gm.

A ***gram-atom*** of an element is a quantity in grammes equal to its atomic weight. Similarly a ***gram-molecule*** of a substance is a quantity in grammes equal to its molecular weight. A ***gram-ion*** is a quantity in grammes of an ion equal to the sum of the atomic weights of its components. Each gram-ion of a monovalent substance carries a charge of 96,450 coulombs. If the valency is Z, the charge on a gram-ion is 96,450 . Z coulombs.

Molecular Conductivity and Equivalent Conductivity.—If χ is the resistivity of a given electrolyte at a fixed temperature, its reciprocal σ is termed the ***conductivity*** of that electrolyte. If χ is expressed in ohm.cm., σ is given in ohm.$^{-1}$cm.$^{-1}$ The ***molecular conductivity***, Σ, is defined by the relation

$$\Sigma = \text{molecular conductivity} = \frac{\sigma}{\text{Number of gram-molecules.cm.}^{-3}}.$$

The ***equivalent conductivity***, Λ, is defined by the relation

$$\Lambda = \text{equivalent conductivity} = \frac{\sigma}{\text{Number of gram-equivalents.cm.}^{-3}}.$$

The unit for equivalent conductivity is obtained as follows:

$$[\Lambda] = \frac{\text{ohm.}^{-1}\text{cm.}^{-1}}{\text{gm.-equiv.cm.}^{-3}} = \text{ohm.}^{-1}\ (\text{gm.-equiv.})^{-1}\text{cm.}^{2}.$$

Further if N = number of gram-molecule.cm.$^{-3}$, and if Z is the valency of the metallic radicle, ZN = number of gram-equivalent.cm.$^{-3}$.

$$\therefore\ \Sigma = \frac{\sigma}{N}, \quad \text{and} \quad \Lambda = \frac{\sigma}{ZN}.$$

$$\therefore\ Z\Lambda = \Sigma.$$

Ionic Mobilities.—The distance through which an ion moves per second when the potential gradient is one volt.cm.$^{-1}$ is termed the ***mobility*** of the ion. A discussion of the methods of determining mobilities would take us beyond the scope of this book so that we shall be content with a description of a direct method of determining the mobility of a Cr_2O_7 ion. This is possible since such ions colour the solution through which they move. Aqueous solutions of potassium bichromate and of potassium carbonate, each having the same resistivity, are arranged as indicated in Fig. 47·10. A and B are two platinum electrodes about 30 cm. apart. The voltage across the tube is such that the potential gradient is about 3 volt.cm.$^{-1}$ This is uniform since the resistance per unit length is constant. The Cr_2O_7 ions move towards A, their motion causing the line of demarcation between the two solutions to travel upwards at a rate of about 1 cm. in 10 minutes under the existing potential gradient.

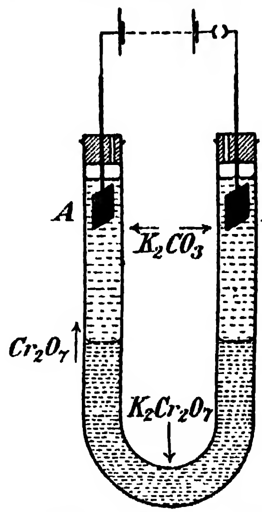


FIG. 47·10.—Ionic Mobilities.

Relative Ionic Mobilities.—Let us consider the motions of the ions of a simple electrolyte. If, under a given potential gradient, the anion and cation move with different velocities, such differences are manifested by the changes in concentration which occur in the electrolyte near to the electrodes. Let Fig. 47·11 represent successive stages in the electrolysis, where the vertical dotted line indicates a permeable diaphragm dividing the cell into two parts. Let the velocity of the positive ions be twice that of the negative ions. Initially the molecules are assumed to be as in the first row and for simplicity we shall suppose that the distance from one molecule to the next is constant—call it λ. At a later instant when the positive

ion has moved a distance 2λ the negative ion will have moved a distance λ and the state of affairs is represented in the second row. We notice that three charges of each sign have been liberated and that in the region of the anode (the anolyte) there has been a

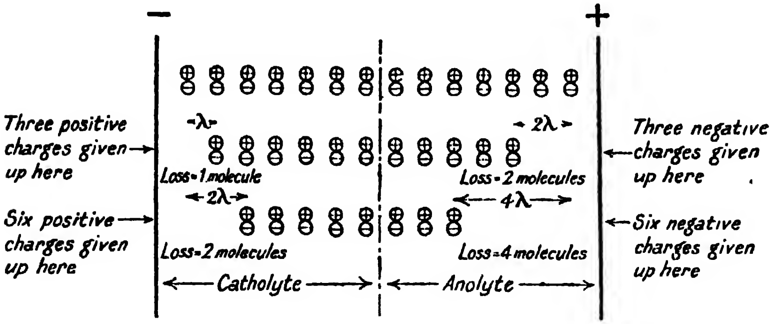


FIG. 47·11.

diminution in concentration of two molecules while the catholyte has suffered a loss of one molecule.

Similarly, in the stage represented in the third row when the positive ion has travelled a distance 4λ and the negative ion a distance 2λ, there is a liberation of six charges of each sign, and a loss of four molecules in the anolyte and two in the catholyte.

This result is perfectly general and we may write

$$\frac{\text{Diminution in concentration at anode}}{\text{Diminution in concentration at cathode}} = \frac{\text{velocity of positive ion, i.e., cation}}{\text{velocity of negative ion, i.e., anion}} = \frac{u_+}{u_-}$$

Hence, by determining the changes in concentration at the anode and cathode we discover the ratio of the mobilities of the particular ions investigated, for the ratio of the mobilities is equal to that of the ionic velocities under the given experimental conditions.

It is more usual to express the results of such investigations in terms of the ***transport numbers*** of the ions. The transport number, n_+, of a positive ion is defined as the ratio

$$n_+ = \frac{u_+}{u_- + u_+}.$$

Similarly
$$n_- = \frac{u_-}{u_- + u_+}$$

Hence $n_+ + n_- = 1$.

The transport number of an ion in a given electrolyte indicates

that fraction of the total current carried by such an ion during electrolysis.

Ionic Velocities and Mobilities.—It has just been shown how the ratio of the velocities of the ions present in a solution, and hence the ratio of their mobilities, may be determined in a given instance. In order to obtain absolute values of these mobilities it is necessary to obtain another relation between them. KOHLRAUSCH first did this by finding an expression for the conductivity of a solution containing the ions in question. Since the conductivity could be measured, the other relation between the mobilities of the ions then became known.

Let us assume that the concentration of the solution is m mole. cm.$^{-3}$; further, let there be complete dissociation, so that the concentration of the ions is also m gram-ions cm.$^{-3}$ if, for simplicity, we assume each gram-molecule to be capable of dissociating into two gram-ions. Each gram-ion of the positive ion has a charge 96,490 coulombs associated with it, if we assume the ion to be monovalent. Similarly, there is an equal amount of negative electricity carried by one gram-equivalent of the negative ion.

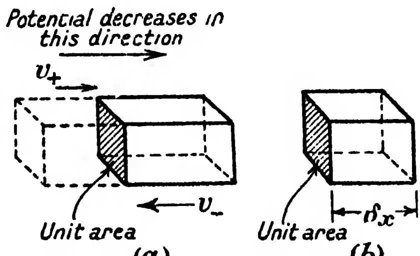


FIG. 47·12.—Ionic Velocities.

Consider a plane of unit area at right angles to the direction of the current—cf. Fig. 47·12 (a). Then in one second the amount of positive electricity passing across this plane is

$$mv_+ \,.\, 96{,}490 \text{ coulomb.},$$

where v_+ is the velocity of the cation.

Similarly, mv_- 96,490 coulombs of ***negative*** electricity pass per second in the ***opposite*** direction. The effective transport of electricity is the sum

$$m(v_+ + v_-) \,.\, 96{,}490 \text{ coulomb.},$$

since there are unlike charges moving in opposite directions.

To obtain another expression for the flow of electricity per second, let σ be the conductivity of the solution (the conductivity is the reciprocal of the resistivity). Suppose that V is the drop in potential per unit length normal to the plane considered. Let δx, Fig. 47·12 (b), be the length of a small element of the solution

of cross-section unity. Then the resistance of this element is given by

$$R = \frac{1}{\sigma} \cdot \frac{\delta x}{1}.$$

Hence the current is given by

$$I = \frac{\text{potential difference}}{\text{resistance}} = \frac{V \cdot \delta x}{\frac{\delta x}{\sigma}} = \sigma V.$$

The quantity of electricity passing per second is therefore σV.

$$\therefore\ (v_+ + v_-) = \frac{\sigma V}{m 96{,}490}.$$

Direct Determination of the E.M.F. of a Cell.—A coil R, Fig. 47·13, whose resistance over a small range of temperatures has been determined, is immersed in oil which is well stirred. A steady current from a battery D is passed through this coil and it is measured by the copper voltameter V. The cell, E, whose e.m.f. is required is connected through a resistance of 10,000 Ω and galvanometer to the ends of R. The current is adjusted until the galvanometer deflexion is zero. The 10,000 Ω resistance serves to prevent large currents being taken from the cell. When an approximate balance has been obtained the resistance S is short-circuited by the key K_2 and the balance point redetermined. If I is the current through V and also through R when the galvanometer deflexion is zero, the potential difference across R is IR where R is the resistance of this coil at the mean temperature of the experiment; this p.d. must be equal to E.

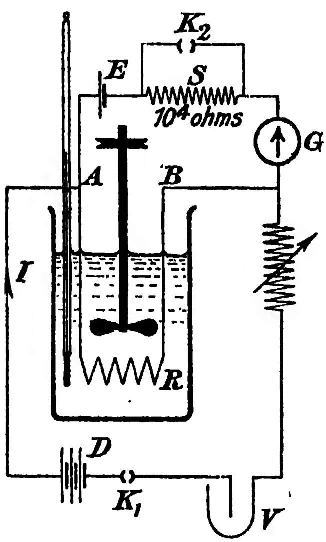


FIG. 47·13.—Direct Determination of e.m.f. of a Cell.

Current Efficiency.—We have seen that the quantity of electricity necessary to produce one gram-equivalent of a substance by electrolysis is 96,490 coulombs or 26·8 ampere-hours. Now this is always the ***minimum*** quantity required. This is not caused by any invalidity of Faraday's laws, for careful experiments carried out both with aqueous and non-aqueous solutions and with fused salts have shown that one gram-equivalent of a substance is

always set free at the electrode when 96,490 coulombs of electricity pass through the electrolyte. Frequently, however, the first product of the electrolysis may react with the electrode or electrolyte or be separated in a form difficult to collect. These are some of the reasons why more electricity than the theoretical quantity is necessary to liberate a specified amount of substance. As an extreme instance we may cite the electrolysis of aqueous potassium chloride between copper electrodes. Although potassium and chlorine are the immediate products of the electrolysis neither appears, for the potassium reacts with the water forming potassium hydroxide and hydrogen, while the chlorine attacks the electrode at which it is liberated and forms cuprous chloride.

The ratio of the yield actually obtained to that calculated from Faraday's laws is termed the ***current efficiency.***

Some Practical Applications.—The process of electrolysis plays an important rôle in many industries. Copper is refined for use in electrical cables, aluminium is obtained from its ores, base metals are coated with more expensive metals, hydrogen and oxygen are prepared electrolytically, chlorine is obtained from sea water, and in recent years heavy water has been prepared in bulk, etc.

Aluminium plays such an important part in modern life that its preparation by an electrolytic process, first described by Héroult in 1886, will be considered briefly. The electrolyte is a solution of alumina (Al_2O_3) in cryolite (Na_3AlF_6) at about 1000° C. The cell,

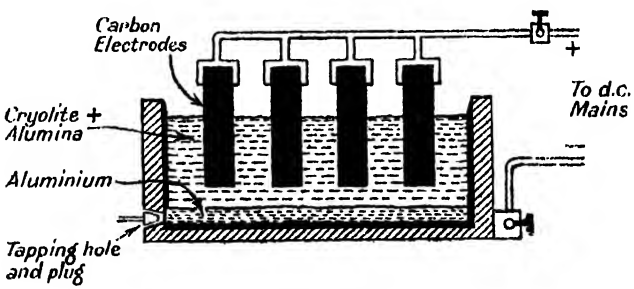


FIG. 47·14.

Fig. 47·14, is an iron box the sides of which are lined with a stamped-in carbon mass, so that a hearth of high electrical conductivity is obtained. The iron box serves as the negative electrode: the anode consists of many rods of pure carbon. The furnace is started by introducing cryolite which is brought to a state of fusion by electrical heating. Pure alumina, prepared from bauxite, is fed in gradually at the surface of the cryolite and as it dissolves the p.d. across the cell falls to about 5 volts: at this stage the current is in the range 15–30 kiloamps. Starting the furnace is

facilitated by a contribution of the molten flux from another furnace.

Initial impurities in the cryolite are soon removed and then if pure alumina only is added pure aluminium is obtained. This collects in the bed of the furnace and is withdrawn at intervals through a tuyère or tapping hole.

The reaction which occurs is equivalent to

$$2Al_2O_3 + 3C \rightarrow 4Al + 3CO_2.$$

In the solution $Al^{\cdot\cdot\cdot}$ and O'' ions are present: in the electrolytic process just described the $Al^{\cdot\cdot\cdot}$ ions pass to the cathode while the O'' ions move to the carbon anode with which, having lost their charges, they react to form carbon dioxide. Some carbon monoxide is produced in increasing quantities as the supply of alumina decreases. The process is continuous, however, for about eight months, for fresh alumina may be added as required. At the end of eight months the cell must be relined with carbon.

In order to silver-plate a copper object, its surface is freed from all trace of grease and suspended, as a cathode, in a solution of silver cyanide ($AgCN$) in aqueous potassium cyanide (KCN).

$$AgCN + 3KCN \rightarrow KAg(CN)_2 + 2KCN.$$

The electrolyte deteriorates gradually—potassium carbonate is formed—so that potassium cyanide must be added. The excess of this substance may help to lower the resistance of the voltameter. A current density of less than 0·003 amp.cm.$^{-2}$ is used. The silver is deposited on the cathode as a dull matt layer which is afterwards polished.

Silver-plating is best carried out at room temperature. For gold-plating a solution of $AuCN$ in KCN is used. Objects of intricate design are best plated at high temperatures because of the higher rate of diffusion of the electrolyte into the interstices of the article.

In the manufacture of electric lamps it is always necessary to test the 'life' of several bulbs selected from a given batch. To do this a copper sulphate solution is placed in series with the lamp. The current is switched on, and continues until the filament of the lamp breaks. The deposition of copper in the electrolytic cell ceases, and from the mass of copper deposited the duration of the current is found, i.e. the 'life' of the bulb is known. Nowadays when lamps can be run for at least 1,000 hours, it is better to shunt the electrolyte with a small resistance, so that only a known fraction of the current through the lamp is available for the electrolysis. This procedure enables smaller quantities of copper sulphate solution to be used.

Electrolytic action is also the cause of much annoyance. If an electric cable passes through a damp region, electrolytic action is

set up, for the water contains dissolved salts, and this in time eats the cable away. When the cable has only one or two strands remaining the electrical resistance is high compared with its original value, so that considerable heat may be developed at this spot. In extreme cases the heat liberated causes a fire.

Secondary Cells or Accumulators.—When one of the primary cells described on pp. 801 to 806 has been in use for some time it ceases to supply current—this process of exhaustion is a gradual one, and the cell cannot be restored to its initial state by the passage of an electric current in the opposite direction. In 1859 PLANTÉ designed a cell, consisting of two lead plates immersed in dilute

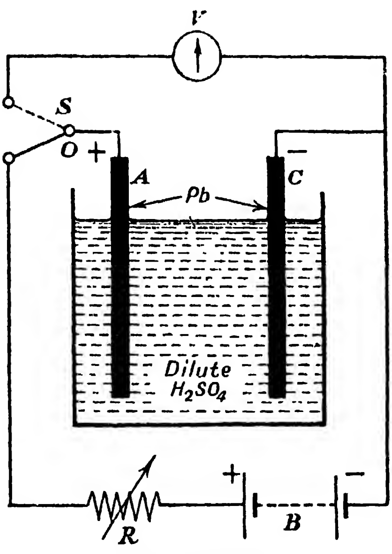


FIG. 47·15 (*a*).

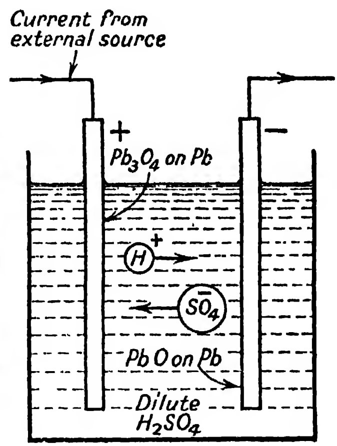


FIG. 47·15 (*b*).

sulphuric acid. Such a cell was found to be almost completely reversible and is termed an ***accumulator*** or ***secondary cell.*** Such a cell must be given electrical energy which is stored as chemical energy: when the plates of the cell are connected to a load, i.e. joined to some piece of electrical apparatus through which current may be passed, this chemical energy is released as electrical energy.

To study the action of such a cell let A and C, Fig. 47·15 (*a*), be two lead plates immersed in dilute sulphuric acid. A battery B, an adjustable resistance R, a two-way switch, and a voltmeter V are connected as shown. With the movable arm of the two-way switch, S, in the appropriate position, a current is passed through

the cell in the direction from A to C. Hydrogen appears at the cathode (−) and oxygen at the anode (+). After a short time the surface of A is covered with a brown layer of lead peroxide (PbO_2); the hydrogen at C does not enter into chemical combination with the lead so that this plate remains in the metallic state. The cell is now in a 'charged' condition, i.e. it is capable of supplying electrical energy. This is easily demonstrated. By altering the position of the movable arm of S the voltmeter V is placed across the cell. At first its reading is about 2 volts, but after a short time the voltage drops rapidly to zero.

During this experiment it may be noticed that when a current is passed through the cell from B, the plate A becomes covered with a chocolate-coloured substance: it is lead peroxide, PbO_2. The plate C is unchanged. When the cell is discharged the lead peroxide disappears and both plates are covered with lead sulphate, $PbSO_4$. The charging process may now be repeated, the oxygen at A converting the $PbSO_4$ into the peroxide, while the hydrogen at C reduces the sulphate to lead. During these processes the lead at C becomes spongy so that a greater surface is available for use, but the mechanical strength of the cell has been impaired.

In order to obviate the tedious process of forming the lead plates, FAURE, in 1880, made the plates with a paste consisting of red lead and sulphuric acid. This is equivalent to a lead sulphate paste, the plates being 'formed' by the passage of a suitable current. Modern accumulators owe their high efficiency to the fact that the plates are now made in the form of a lead-grid, into whose openings certain 'pastes' are pressed. The grids secure the pastes more effectively than in the original method. For the plate which is to serve as the positive electrode of the accumulator, the paste is red lead, Pb_3O_4, the 'binding material' being sulphuric acid. The openings of the other plate contain litharge, PbO, with sulphuric acid as the binder. These pasted plates are not ready for use but have to be processed so that the litharge is converted to spongy lead and the red lead to lead peroxide.

When a potential difference is applied to the pasted plates, immersed in dilute sulphuric acid—cf. Fig. 47·15—the sulphions (SO_4'') move towards the anode, lose their charges, and the following reactions occur:—

$$2SO_4 + 2H_2O \rightarrow 2H_2SO_4 + O_2.$$

$$Pb_3O_4 + O_2 \rightarrow 3PbO_2.$$

At the same time the hydrogen ions ($H^{\cdot}$) move to the cathode where the litharge is reduced to spongy lead: the action is given by

$$PbO + H_2 \rightarrow Pb + H_2O.$$

The Discharging and Re-charging of an Accumulator.—When the accumulator is sending current through an external circuit as in Fig. 47·16 (*a*) the e.m.f. causes the hydrogen ions ($H^{\cdot}$)

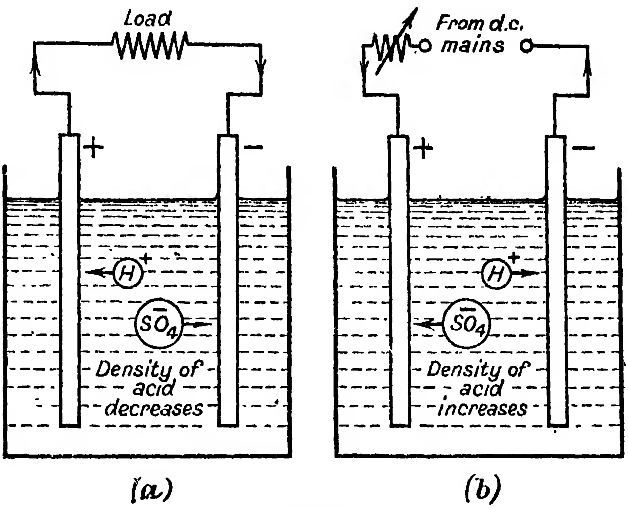


FIG. 47·16.

to move to the anode and the sulphions (SO_4'') to the cathode. The reactions which occur are

(*a*) At the negative electrode.

$$Pb + SO_4'' \rightarrow PbSO_4 + 2\,\varepsilon$$

where the symbol ε denotes an electron. These electrons pass to the positive electrode through the external circuit.

(*b*) At the positive electrode.

$$PbO_2 + 2H^{\cdot} + H_2SO_4 + 2\,\varepsilon \rightarrow PbSO_4 + 2H_2O.$$

To recharge the cell a current must be passed through it—cf. Fig. 47·16 (*b*). As when the plates are 'formed' the hydrogen ions ($H^{\cdot}$) travel through the electrolyte to the cathode and the sulphions (SO_4'') to the anode. The following reactions occur :—

(*a*) At the anode.

$$PbSO_4 + SO_4'' + 2H_2O \rightarrow PbO_2 + 2H_2SO_4 + 2\,\varepsilon.$$

The electrons which are liberated pass through the electrolyte.

(*b*) At the cathode.

$$PbSO_4 + 2H^{\cdot} + 2\,\varepsilon \rightarrow Pb + H_2SO_4.$$

These equations show that during the process of charging the

cell sulphuric acid is set free, i.e. the density of the electrolyte increases; during discharge the density falls.

The state of a cell is ascertained by observing the density of the acid. The acid[1] of a fully charged cell has a density of 1·25 gm. cm.$^{-3}$, at room temperature.

The curves in Fig. 47·17 indicate the way in which the e.m.f.

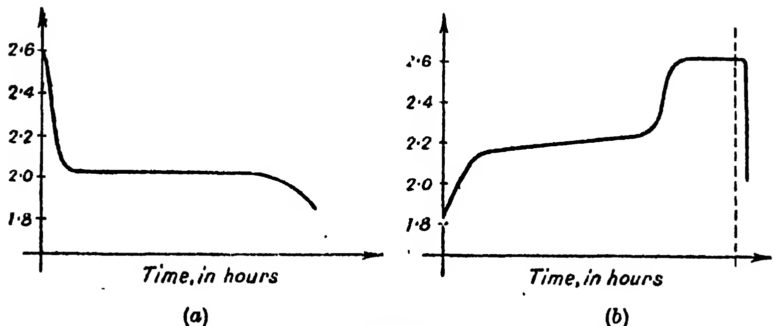


FIG. 47·17.

of the cell (as measured by a high resistance voltmeter placed across its terminals) varies with time during the process of (*a*) discharging, (*b*) charging.

The use of accumulators is due to the fact that their e.m.f. is large, 2 volts, and their internal resistance is low, so that they can supply large currents. Unlike primary cells they may be recharged. Against these assets must be set the following disadvantages; their cost is high, they must be treated carefully, and their mass is considerable. After about two or three years' use their efficiency is very low, i.e. only a small fraction of the electrical energy spent in charging them is re-available. It is believed that this gradual decline in the efficiency is due to traces of iron in the lead plates.

How to Charge Lead Accumulators.—Let T_1 and T_2, Fig. 47·18, be the terminals of the mains supplying direct current. It will be assumed that the negative terminal is earthed and that the main switch is on the high potential side of the installation. Suppose that B is the battery to be charged. The negative terminal of B is connected to T_1 through an ammeter A. L_1, L_2, L_3, etc., are lamps in parallel with one another. Switches K_1, K_2, K_3, etc., are arranged so that the number of lamps in parallel may be varied and the current adjusted to lie within the limits of the charging current specified by the manufacturers of the cells.

A battery is fully charged when, with the current flowing at the normal charging rate, all cells are gassing freely and evenly, and

[1] Dilute H_2SO_4 of this strength may be made as follows: add 289 cm.^{3} conc. H_2SO_4 to 1,000 cm.^{3} distilled water.

the density of the acid is a maximum, viz. 1·25 gm.cm.$^{-3}$ at normal temperatures.

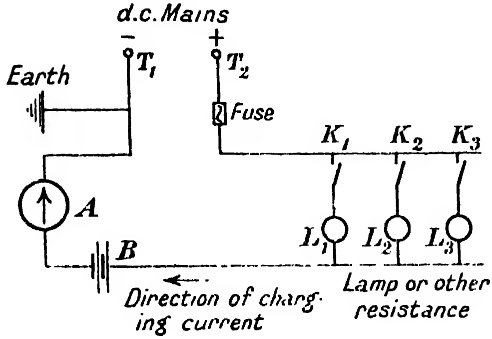


FIG. 47·18.—Charging Accumulators.

On the Care of Accumulators.—If the cells are received 'dry,' i.e. without containing acid, they should be filled with sulphuric acid of density 1·25 gm.cm.$^{-3}$ to the 'acid-level' line, i.e. to a height about 1 cm. above the top of the plates. The cells should be allowed to stand for twelve hours and sufficient of the above acid then added to restore the acid to its original level. The battery should then be charged at its normal rate for two days, the temperature never being allowed to rise above 40° C. During the end stages of this charging process, gas should be freely evolved from the cells and the voltage across each cell should remain constant.

The cell is then ready for use. The state of its charge at subsequent times is ascertained from observations on the density of the acid by means of a hydrometer. It has already been mentioned that the cell is fully charged when the density of the acid is 1·25 gm.cm.$^{-3}$: at half charge the density is 1·18 gm.cm.$^{-3}$: the cell is fully discharged when the density is 1·11 gm.cm.$^{-3}$. Under no conditions should the acid density be allowed to fall below 1·16 gm.cm.$^{-3}$ so that the formation of lead sulphate on the plates—'sulphating'—is thereby reduced.

From time to time distilled water must be added to the cell to compensate for evaporation. If a cell is to remain idle it should be fully charged and then 'refreshed' every month.

Great care should be taken never to short-circuit an accumulator since when a heavy current is taken from the cell its plates tend to become buckled. The heat evolved during such a discharge tends to loosen the material on the plates.

It is also essential to see that the acid used is free from dissolved metallic salts since the metals would eventually be deposited on the plates by electrolysis and local action occur.

The Edison Storage Cell.—In recent years a new form of accumulator has appeared; it is known as the Edison storage cell. It was designed with a view to being less massive but more robust than the lead accumulator. The positive plate is composed of nickel hydroxide and flake nickel or graphite; the nickel hydroxide constitutes the active material, but since it is a non-conductor of electricity the

nickel or graphite is added to render the plate conducting. After the plate has been 'formed' the nickel hydroxide is replaced by nickel peroxide—this latter substance is used in the presence of water, so that its formula may be written $Ni(OH)_4$, i.e. $NiO_2 + 2H_2O$. The active material in the negative plate is finely divided iron. These two plates are immersed in a 20 per cent. solution of caustic potash containing a small amount of lithium chloride.

The active mixture for the positive plate is compressed into a steel tube which is perforated over its cylindrical surface so that the alkali may have easy access to it. The tubes are made of very thin cold-rolled carbon steel which is nickel-plated after the tube has been made. The tubes are about 12 cm. long and 0·5 cm. in diameter.

The negative plate consists of steel boxes each 7·5 cm. long, 1·2 cm. wide and 0·3 cm. thick. They are made of steel and are nickel-plated. The finely divided pure iron is placed in these pockets, and a trace of mercuric oxide (HgO) added to lower the resistance of the electrode.

The chemical reactions which occur in the cell may be summarized by means of the following equation :—

$$Fe + KOH + 2Ni(OH)_4 \underset{\text{charge}}{\overset{\text{discharge}}{\rightleftharpoons}} 2Ni(OH)_2 + KOH + Fe(OH)_2.$$

It will be noticed that the caustic potash does not vary in amount, so that the density of the solution is no indication of the state of the cell. The e.m.f. of the Edison storage cell is 1·35 volts.

To Determine the Polarity of a Cell or the Mains.—An aqueous solution of potassium iodide is prepared, and a small quantity of starch paste is added. A piece of filter paper is moistened with this solution. If now two wires joined to the electrodes of a cell are allowed to touch this paper, a feeble current passes through the solution which is on the paper. Iodine is liberated at the anode which, acting on the starch, produces a dark blue compound. If the mains are under test a lamp must be placed in series with them. If starch paper is not available the electrodes may be dipped into salt water (in an egg-cup). Bubbles of hydrogen appear at the cathode. Alternatively, if a small quantity of phenolphthalein is added to the sodium chloride solution, a pink colour appears at the cathode: this colour is due to the formation of sodium hydroxide.

Similarly, red litmus paper moistened with sodium chloride solution turns blue at the cathode.

Depolarizers.—In an earlier chapter [cf. p. 801] it has been shown how the polarization may be prevented by the use of a suitably chosen depolarizing agent. In addition to the inorganic salts which have been mentioned it is found that many organic compounds can also be similarly used. ***Cathodic depolarizers*** are those substances which either take up hydrogen or yield oxygen, or else do both simultaneously, i.e. cathodic depolarizers are reduced. ***Anodic depolarizers*** are substances which can be oxidized.

Electrolytic Reduction.—An electrochemical method of preparing an organic compound possesses several advantages over a purely chemical method, for, by varying the conditions under which the depolarizing agent is reduced, it is possible to prepare a series of compounds from one such depolarizer. The electrolytic reduction of a depolarizer proceeds in two stages.

(*a*) The positively charged hydrogen atoms lose their charges and become atomic hydrogen. $H^{\cdot} + \ominus = H$.

(*b*) The atomic hydrogen combines with the depolarizing agent and reduces it.

By such means as this aniline ($C_6H_5NH_2$) is prepared from nitrobenzene ($C_6H_5NO_2$); indigo is formed from indigo white; oleic acid is converted into stearic acid.

Electrolytic Oxidation.—In this process the negatively charged hydroxyl ions lose their charges and oxygen is liberated—

$$2OH' + 2\oplus = H_2O + O.$$

For commercial purposes electrolytic oxidation is used in the preparation of formic acid (H.COOH) from acetylene (C_2H_2)—

$$C_2H_2 + 6OH' + 6\oplus = 2H.COOH + 2H_2O.$$

When an aqueous solution of alkali halide containing acetone $CH_3.CO.CH_3$ or alcohol $CH_3.CH_2.OH$ is electrolysed, chloroform, bromoform or iodoform is formed. The first stage of the process consists in the formation of ClO′ ions [or the corresponding ions with bromine or iodine].

$$2Cl' + 2\oplus = 2Cl.$$
$$2Cl + 2OH' = ClO' + Cl' + H_2O.$$

The ClO′ then reacts with the acetone as follows:—

$$CH_3.CO.CH_3 + 3ClO' = CHCl_3 + CH_3CO_2' + 2OH'$$

In the case of alcohol the equation is

$$CH_3.CH_2.OH + 5ClO' = CHCl_3 + H_2O + 2Cl' + 3OH' + CO_2.$$

EXAMPLES XLVII

1.—A tangent galvanometer has 6 turns of wire of mean radius 10 cm. How much copper will be deposited in 30 mins. by a current which produces a deflexion of 45°? [e.c.e. for copper = 0·000328 gm. coulomb^{-1}, H_0 = 0·18 oersted.]

2.—When a current is passed through a voltameter and tangent galvanometer, a deposit of 0·32 gm. of copper is obtained per hour and the deflexion is 39°. The diameter of the coils is 18 cm. Assuming that the above numbers are liable to experimental error, calculate the number of turns in the coil. [H_0 = 0·18 oersted.]

3.—A piece of very thin metal measures 8 cm. × 18 cm. It is desired to coat it with a layer of silver 0·1 mm. thick. For how long must a current of 3·5 amperes be passed. [e.c.e. for silver = 0·001118 gm. coulomb^{-1}; density of silver = 10·5 gm. cm.$^{-3}$]

4.—State Faraday's laws of electrolysis. How may they be verified?

5.—Define the terms ampere, electrolyte, cation, ohm, electron.

6.—What is meant by the statement that the back e.m.f. in an electrolyte is 0·2 volts? A battery having a total e.m.f. of 20 volts and 1 ohm internal resistance is connected in series with an electrolyte. This is shunted with a 10 ohm coil. If the battery supplies a current of 3 amperes and the back e.m.f. in the electrolyte is 0·1 volt deduce the resistance of the electrolyte.

7.—Explain what is meant by the statement 'the electrochemical equivalent of copper is 0·000329 gm. coulomb^{-1}.' Calculate the cur-

rent through a copper voltameter if 0·987 gm. of copper is deposited in it in 40 minutes.

8.—Explain Ohm's law and describe how you would verify it. Discuss whether the law holds for electrolytic conductors.

9.—Give a short account of the laws of electrolysis. A Daniell cell is used to send a steady current through a certain circuit. It is found that in half an hour the negative pole of the cell has decreased in weight by 0·070 gm. Calculate the increase in weight of the positive pole, and the mean value of the current supplied by the cell. [The atomic weights of copper and of zinc may be taken as 63·6 and 65·4 respectively, and the electro-chemical equivalent of hydrogen as 0·0000104 gm. coulomb.$^{-1}$]

10.—State Faraday's laws of electrolysis and describe how you would proceed to measure the electrochemical equivalent of silver.

11.—Describe and explain what happens when an aqueous solution of copper sulphate is electrolysed between (*a*) soluble electrodes, (*b*) insoluble electrodes.

12.—Give an account of the conduction of electricity through aqueous solutions of inorganic salts. How does it differ from the conduction of electricity through mercury ?

13.—Describe how you would investigate the validity of Ohm's law in the case of acidulated water and state the results you would expect to obtain.

14.—Define the ampere and the electromagnetic unit of current. Explain how with the aid of a tangent galvanometer and a copper voltameter you would determine the relation between these two units. —(L. '29.)

15.—Explain with the aid of a circuit diagram how an electrometer may be used in an experiment either to find the resistance of an electrolyte, or to compare the capacities of two condensers. (N.H.S.C. '29.)

16.—State Faraday's laws of electrolysis.

Describe a lead accumulator and give an outline of the changes that probably occur during the process of charging the cell.

In the electrolysis of acidulated water it is found that 0·116 cm.^{3} of hydrogen (at S.T.P.) is liberated per coulomb. Assuming that the charge carried by a hydrogen ion is $1·60 \times 10^{-19}$ coulomb, calculate the number of molecules in 1 cm.^{3} of hydrogen (at S.T.P.).

17.—If a steady current of 0·25 amp. for 0·5 hr. liberates by electrolysis 52 cm.^{3} of hydrogen at S.T.P. and 1 mole. contains $6·03 \times 10^{23}$ molecules, calculate a value for the charge on an hydrogen ion in electrolysis.

CHAPTER XLVIII

ELECTROTHERMAL EFFECTS AND THERMOELECTRICITY

Electrothermal Effects

Joule's Original Experiment on the Production of Heat by Current Electricity.—In this research Joule proposed to investigate how the amount of heat dissipated by a current flowing in a conductor varied, in a given time, with the strength of the current. The apparatus, shown in Fig. 48·1, consisted of a tall glass cylinder containing water. The wire was passed through a

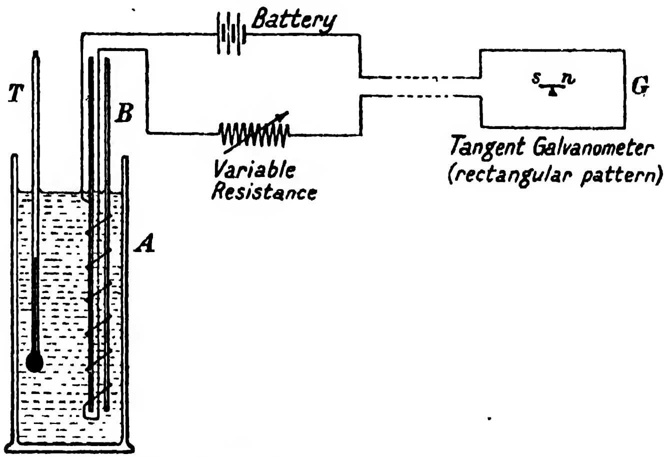


Fig. 48·1.—Joule's first Experiment on the Heating Effect of a Current.

thin glass tube B and then closed upon it. The extremities of the coil were then pulled slightly apart so that the wire was not short circuited. In some experiments a cotton thread was interposed between the windings. The current was measured by a type of tangent galvanometer, G, the frame being rectangular; its indications were standardized by a voltameter—Joule expressed his quantities of electricity (current × time) 'on the basis of Faraday's great discovery of definite electrolysis.' Joule used a mercury-in-glass thermometer, the scale being graduated on the

stem. Before making an observation on the temperature, the water was stirred by means of a feather. Joule took the precaution of using the thermometer in a vertical position and to bring his eye 'to a level with the top of the mercury.' He estimated changes in temperature of 0·1° F.

As a result of these experiments Joule found that the heat generated in a given time in a specified conductor is directly proportional to the square of the current; also, for a given current, the heat, Q, produced in a given time was directly proportional to the resistance of the wire, i.e.

$$Q \propto I^2R.$$

The above work was carried out at a time when Joule was not acquainted with Ohm's law. In fact, if Joule had been aware of the validity of Ohm's law, he could have deduced the above result theoretically as we shall see below. In fact, Joule's law and Ohm's law are really alternatives.

The Heating Effect of a Current.—From the definition of the electromagnetic unit of potential difference it follows that when a current i [e.m.u.] is flowing between two points differing in potential by an amount e [e.m.u.] the energy, W, liberated in t seconds is eit ergs, since it is the quantity of electricity transferred in this time. In practice it is found that inconveniently large numbers occur when the electromagnetic system of units is employed. When the current is measured in amperes (I) and the potential difference in volts (V) [practical units] the work (W) liberated in t seconds is measured in joules and we have $W = VIt$ joules $= VIt \times 10^7$ ergs. The validity of the above is established when we remember that one e.m.u. of current $= 10$ amperes, and one e.m.u. of potential $= 10^{-8}$ volts, for then $W = eit = V \times 10^8 . I \times 10^{-1} . t = VI \times 10^7 . t$ ergs, or VIt joules. Now the work done per second is termed the ***power***, the practical unit of which is the ***watt***. Hence the power necessary to send a current I amperes through a conductor across which the p.d. is V volts is VI watts.

If R is the resistance in ohms of the above circuit, $R = \frac{V}{I}$, so that $W = I^2Rt$, or $\frac{V^2t}{R}$, joules.

The above equations give us the energy used in 'overcoming the ohmic resistance of the conductor'. All this energy is dissipated as heat in the conductor and the amount of heat, Q, generated in t seconds is given by $Q = \frac{W}{J}$, where J is the mechanical equivalent of

heat, i.e. $Q = \frac{VIt}{J} = \frac{VIt}{4{\cdot}18}$ cal. Sometimes this is written $Q = \frac{V^2t}{4{\cdot}18R}$, but in precision work it is better to measure V and I since the resistance of the conductor depends upon its temperature and, in general, this resistance would be measured at room temperature and this is certainly not the temperature of the wire when heat is being developed in it.

The formula just obtained may be derived in a slightly different way as follows. Let A and B be two points in a wire at potentials V_1 and V_2 respectively, $(V_1 > V_2)$. Suppose that a charge δq passes from A to B. At A the potential energy associated with the charge is $V_1 . \delta q$ [the potential energy in such a case is the work done in bringing up the charge δq from infinity (zero potential) to A]. At B the potential energy of this charge is $V_2 . \delta q$. Hence the loss in potential energy associated with this charge when it moves from A to B is $(V_1 - V_2)\delta q = (V_1 - V_2)I . \delta t$, if I is the current flowing for a time δt. This energy appears as heat. The rate at which energy is dissipated is therefore $(V_1 - V_2)I = VI$, if $(V_1 - V_2) = V$, as before. If V is in volts and I in amperes, the product VI is in watts.

Verification of Joule's Law.—The apparatus consists essentially of a calorimeter containing paraffin oil or aniline [low vapour pressure to diminish evaporation and consequent heat loss, and a small specific heat so that the temperature rises quickly]. This calorimeter is fitted with an ebonite lid which supports the heating coil—see Fig. 48·2. The current is supplied by means of a battery and is regulated, i.e. it is kept constant, with the aid of the sliding resistance S. An ammeter A indicates the magnitude of the current. The potential difference across the coil is measured by the voltmeter V which is in parallel with the heating element [or a potentiometer may be used]. The heat developed in a given time t secs. may be calculated as follows when the rise in temperature of the calorimeter and its contents has been determined.

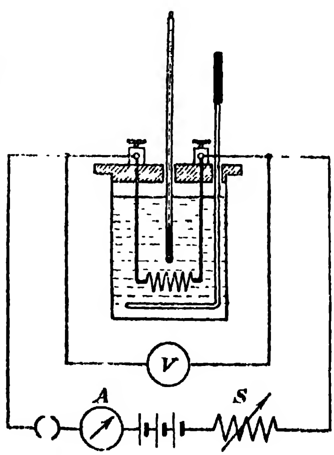


FIG. 48·2.—Verification of Joule's Law of Heating (or Determination of J (electrical).

Let M be the mass of the oil of specific heat σ, m the mass

of the calorimeter whose material has a specific heat s, and θ the observed rise in temperature. Then the heat developed is $(M\sigma + ms)\ \theta$ cal.; this should be equal to $\dfrac{VIt}{4 \cdot 18}$ cal.

Better results will be obtained if the calorimeter is surrounded by a double-walled vessel containing water but insulated thermally from it as indicated in Fig. 10·4 (*a*). This protects the calorimeter from draughts and thereby prevents erratic exchanges of heat between the calorimeter and its surroundings. The temperature of the water in this outer vessel should remain constant and be equal to the mean of the initial and final temperatures of the oil, for then heat is gained by the calorimeter from its surroundings during the first half of the experiment and an equal amount lost to them during the second half. Alternatively, Ferry's method for correcting for the heat exchange between the calorimeter and its surroundings may be used [cf. p. 208].

Definition of Electromotive Force.—Suppose that a generator of electrical energy is connected to a resistance R and that I is the current supplied. Let V be the potential difference across R. Then VI is the rate at which energy is dissipated in R. The total rate at which energy is supplied will be VI + a certain quantity proportional to I, say βI. Then V + β is the rate at which energy is supplied per unit current by the generator. This is termed E, the electromotive force of the generator.

The Laws of Heating.—When a current is passed through a wire the rise in temperature of the wire depends on the material of the wire and the nature of its surface. It is possible to find a very thin wire whose resistance is the same as that of a longer piece of thick wire: the wires may be of the same or of different materials. If their surfaces have the same radiating powers and the same current is sent through them, then an equal number of calories will be developed in each in a given time. Since the mass of the thin

FIG. 48·3.—The Laws of Heating.

wire is small its rise in temperature will be much greater than that of the thick wire. In addition, the thick wire has a larger surface from which heat is radiated so that this tends further to diminish the rise in temperature. The ultimate temperature attained by such wires is reached when the rate at which heat is being developed in them is equal to that at which it is lost from its surface to the surrounding air.

The following experiments provide us with excellent illustrations of the laws governing the heating effects of currents. Fig. 48·3 (*a*) represents a chain formed of alternate links of iron and copper, each of No. 28 S.W.G. If a current of 4 amperes is sent through the composite chain the iron links become red hot and even if the current is increased until the iron wires are burnt out the copper does not emit any visible radiation. The explanation of this is to be found in the fact that the resistivity of iron is about seven times that of copper and since the specific heats of iron and copper are approximately equal the rise in temperature is about seven times as great in the iron wire.

The second experiment consists in placing pieces of copper and iron of equal length and gauge [No. 42 S.W.G.] in parallel and introducing them into a circuit carrying a current of about 10 amperes [cf. Fig. 48·3 (*b*)]. In this instance it is the copper which glows. This is because the heat developed in a resistance at constant voltage is inversely proportional to the resistance. Hence in this instance seven times as much heat is developed in the copper as in the iron.

Maximum Power to be obtained from a Given Battery.—Let E be the e.m.f. of a battery with internal resistance B. If a resistance R is connected to the battery the current through R is $\frac{E}{(B+R)}$. Hence the energy dissipated per second in R is $\frac{E^2R}{(B+R)^2} = P$, say. For this to be a maximum or a minimum $\frac{dP}{dR}$ must be zero. Differentiating we have

$$\frac{dP}{dR} = \frac{E^2(B+R)^2 - E^2R[2(B+R)]}{(B+R)^4}.$$

This is zero when $B = R$. A second differentiation shows that W is a maximum when $B = R$. This important proposition is illustrated by Ex. 7, p. 935.

The Principle of Least Heat.—Let us consider two resistors R_1 and R_2 in parallel, the current through the combination being I. If I_1 is the current through R_1, $(I - I_1)$ is the current through R_2.

$\therefore$ *Rate at which heat is produced in the two resistors*

$$= \frac{I_1^2R_1 + (I - I_1)^2R_2}{J} \text{ cal.sec.}^{-1} \quad \text{[Practical units assumed.]}$$

$$= P \text{ (say)}.$$

$$\therefore \frac{dP}{dI_1} = \frac{1}{J}[2I_1R_1 + 2(I - I_1)(-R_2)].$$

This is zero if $I_1(R_1 + R_2) = IR_2$, and since $\frac{d^2P}{dI_1^2}$ is positive for this value of I_1, the rate of production is a minimum when this condition is satisfied. Thus P is a minimum when the current through each resistor is inversely proportional to its resistance. This, however, is the manner in which the current I does distribute itself and hence the division of current is such that the rate of production of heat is a minimum.

The Rise in Temperature of a Wire due to the Passage of an Electric Current.—Consider a wire of length l, radius r, through which a current I is flowing. If R is the resistance of the wire when steady conditions have been attained, i.e. when ϕ, its temperature excess above its surroundings, is constant, the rate at which heat is developed in the wire is $\frac{I^2R}{J}$. If Newton's law of cooling is assumed to be true,

$$\frac{I^2R}{J} = \varepsilon . 2\pi rl . \phi, \qquad \text{[cf. p. 319 and p. 342]}$$

where ε is the surface emissivity of the wire. But $R = \chi\frac{l}{\pi r^2}$, where χ is the resistivity of the material of the wire at the excess temperature ϕ.

$$\therefore\ I = \pi\left[\frac{2\varepsilon J\phi}{\chi}r^3\right]^{\frac{1}{2}}.$$

Thus the elevation of temperature is independent of the length of the wire, but for a constant current $\frac{\phi}{\chi}$ is inversely proportional to the cube of the radius.

Instead of assuming Newton's Law of Cooling it may be supposed that the wire behaves as a black body when

$$\frac{I^2R}{J} = 2\pi rl\sigma(T^4 - T_0^4),$$

where σ is Stefan's constant, T and T_0 the absolute temperatures of the wire and its surroundings. If the wire is a filament in a lamp $T \gg T_0$, so that T_0^4 may be neglected in comparison with T^4. Since $R = \frac{\chi l}{\pi r^2}$,

$$I^2 = 2\pi^2 Jr^3\frac{\sigma}{\chi}T^4.$$

For tungsten $\frac{\chi}{T}$ is almost constant, so that for wires of constant radius, r,

$$I^2 \propto T^3, \text{ or } T \propto I^{\frac{2}{3}}.$$

Continuous Flow Calorimetry.—The heating effect of a current provides us with an accurate means of determining the specific heat of a liquid. The apparatus shown in Fig. 48·2 may be used for this purpose. The more accurate electrical methods of determining the specific heats of liquids and gases have already been discussed and should be revised at this stage.

A Few Practical Applications.—The heating effects of currents have numerous applications in everyday life; foremost among these are the incandescent electric lamps, radiators, cooking ovens, furnaces, and wireless valves.

The incandescent lamp consists of a glass bulb in which there is a carbon or tungsten filament, the temperature of which is raised when a suitable current is sent through it. It is at once obvious that the higher the temperature of the wire the greater its power as a source of intense light. In order to attain this high temperature a wire of high melting-point must be used, e.g. tungsten or tantalum. At these high temperatures carbon or any metal rapidly oxidizes in the presence of oxygen, so that manufacturers exhaust the lamps. As the temperature at which a lamp was used became higher it was found that the metallic filaments began to evaporate; consequently their resistance increased, the temperature rose and the wire melted: in addition, the evaporated metal was deposited on the glass walls so that the brilliancy of the lamp was impaired. To obviate these disadvantages modern bulbs are filled with nitrogen, the pressure of which is sufficient to render the evaporation losses negligible. In the smaller lamps a small amount of argon is added.

Electric radiators consist of nichrome wire wound on a fireclay support; the resistance of the wire is such that a temperature of about 800° C. is easily obtained when an appropriate current is passed along the wire.

In small electric furnaces, such as are used in a laboratory, nichrome wire is wound on a silica tube and then covered with 'purimachos'—a fireclay cement. The whole tube is supported along the axis of a cylindrical container, the intervening space being packed with asbestos wool. In the very latest type of such a furnace the wire is of molybdenum, and the whole is placed in an atmosphere of 'cracked' ammonia, i.e. ammonia which has been passed through a red-hot tube so that it is dissociated completely. In such an atmosphere molybdenum does not 'burn out' even at a temperature of 1,800° C., e.g. steel may be very easily melted.

In wireless valves and X-ray tubes a tungsten wire is heated to over 2,000° C. when it emits a copious supply of electrons—the 'atoms of electricity.' Under a suitable electric field these can be made to move; they then constitute a current [cf. p. 797].

The arc lamp consists essentially of two carbon rods, generally

at right angles to one another. They are connected to the mains through a suitable resistance, and their distance apart can be adjusted by means of a screw. When the carbons touch, there is a large current which is sufficient to vaporize the carbon at the points where contact is made. The carbons are then separated by a few millimetres, and the current continues to flow and a temperature of over 3,000° C. is reached. Modern steel works use electric furnaces in which very large carbon electrodes are employed. A current of 2,000 amperes at 65 volts is then sufficient to melt two tons of steel and alloys.

In surgery a thin platinum wire heated to redness is often used to cut tissue when, if the temperature is suitable, hæmorrhage is reduced to a minimum.

Hot-wire Instruments.—For some purposes, currents are employed whose direction is reversed many times per second. These are known as alternating currents [cf. p. 953] and it is at once apparent that such currents cannot be measured by any of the arrangements hitherto described since the current is reversed before the moving part of the instrument has changed its zero position. Such currents are usually measured by the heating effects they produce since these are independent of the direction in which

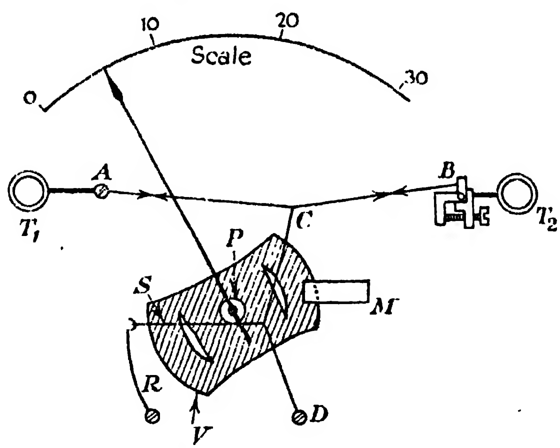


FIG. 48·4.—Hot-wire Ammeter.

the current is flowing. A hot-wire ammeter of the double-sag type is indicated in Fig. 48·4. AB is the hot wire, supported at both ends, and consists of a piece of platinum-iridium wire about 0·1 mm. in diameter. The above material is chosen since it does not oxidize readily and therefore does not deteriorate when subject to high temperatures. A second thin wire, CD, made of phosphor-bronze, is attached to the centre of AB: to this second wire a fine silk

thread S is attached. This passes round a small pulley P and then is attached to a spring, R, so that each portion of the system is kept in tension. A light aluminium pointer is attached to P. When the hot wire expands and the system is made taut by R, the pointer is deflected. Connexions are made to the terminals T_1 and T_2. The scale readings may be checked by passing known direct currents through the instrument. V is an aluminium vane attached to the pulley and it moves between the poles of a horse-shoe magnet M : in this way oscillation is prevented and, moreover, sudden changes in the current do not reveal themselves so quickly and consequently the wire is never subject to excessive stress.

Since in an alternating current the current varies periodically it is obvious that a hot-wire instrument does not measure the current at any instant in the cycle : what it does measure is the ***effective, virtual,*** or ***root-mean-square value*** of the current. This is defined as that steady current which produces the same heating effect per unit time as the alternating current.

The Measurement of Power.—The power or rate at which energy is being dissipated in any portion of an electrical circuit is equal to the product VI watts, where V is the voltage across the portion of the circuit and I is the current in amperes. Instead of measuring the voltage and current separately and deducing the number of watts from them, wattmeters have been designed to measure the power directly. The essential features of such an instrument are indicated in Fig. 48·5. In a wattmeter there are two concentric coils, one being fixed while the other is movable. The fixed coil consists of a few turns of thick wire which are connected to the terminals AA. These enable the fixed coil to be placed in series with the current in the circuit where the consumption of energy is being determined. The movable coil consists of many turns of very thin wire and it is connected to the terminals VV of the instrument. These permit the movable coil to be connected across the mains. The movable coil is wound on a bobbin carrying a pointer. The coils are arranged so that in their zero position

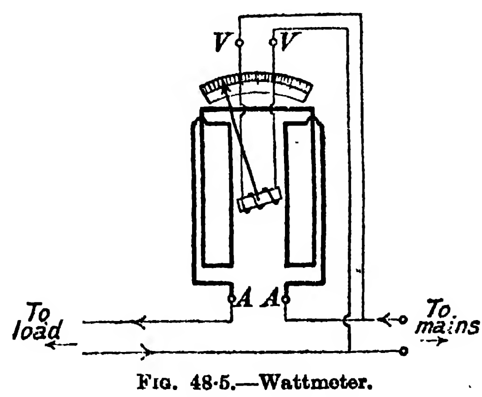


FIG. 48·5.—Wattmeter.

their planes are mutually perpendicular ; hence when currents

pass through them there is a couple tending to make the coils coplanar. The deflexion of a pointer attached to the movable coil is proportional to the rate at which energy is being consumed. The instrument may be calibrated with the aid of a standard voltmeter placed in parallel with VV and a standard ammeter placed in series with AA.

Note on the Measurement of Electrical Energy.—It has already been shown that the power or rate at which energy is produced in a load is VI watts where V is the p.d. across the load in volts, and I amp. is the current through it. Engineers use a larger unit of power : it is the ***kilowatt*** which is equal to 1,000 watts. Another unit of power is the ***horse-power*** which is the power when energy is developed at a rate of 550 ft.lb.-wt.sec.$^{-1}$

$$\begin{aligned}\therefore 1\text{ H.P.} &= 550 \times 30{\cdot}48 \times 453{\cdot}6 \times 980 \text{ erg.sec.}^{-1}\\ &= 746 \times 10^7 \text{ erg.sec.}^{-1}\\ &= 746 \text{ watt.}\end{aligned}$$

Electrical motors and other appliances have their power stamped upon them so that if the time for which they are in use is known the energy consumed may be calculated, since

$$\text{Energy} = \text{power} \times \text{time in seconds.}$$

Electric supply companies measure the energy supplied in kilowatt-hours or Board of Trade units [1 KVA-hour].

$$\begin{aligned}\text{Now 1 KVA-hour} &= 1 \text{ kilowatt-hour}\\ &= 10^3 \text{ joule.sec.}^{-1} \times 1 \text{ hour}\\ &= 10^3 \text{ joule.sec.}^{-1} \times 3600 \text{ sec.}\\ &= 3{\cdot}6 \times 10^6 \text{ joule.}\end{aligned}$$

Example.—A 2,000 watt heater operated from 240 V. mains raises the temperature of 10 litres of water 30° C. If 1 B.T.U. costs 2*d.*, find minimum values for the cost and time required.

$$\text{Heat required} = 10 \times 1{,}000 \times 30 \text{ cal.} = 3 \times 10^5 \text{ cal.}$$

$$\text{Rate of supply of heat by the heater} = \frac{2 \times 10^3}{4{\cdot}18} \text{ cal.sec.}^{-1}$$

$$\therefore \text{ Time} = \frac{3 \times 10^5 \times 4{\cdot}18}{2 \times 10^3} \text{ sec.} \simeq 10\tfrac{1}{2} \text{ minutes.}$$

$$\begin{aligned}\text{Energy consumed} &= 2{,}000 \text{ joule.sec.}^{-1} \text{ for } 630 \text{ sec.}\\ &= 2 \text{ KVA} \times \left(\frac{630}{3600}\right) \text{hr.}\\ &= 2 \times \frac{63}{360} \text{ B.T.U.}\end{aligned}$$

$$\therefore \text{ Cost} = \frac{63}{90}d. = 0{\cdot}7d.$$

Experimental Determination of the Thermal Emissivity of a Wire.—A nickel wire [S.W.G.34] and 30 cm. long is suitable. AB, Fig. 48·6, is the wire whose emissivity is required. It is held in a

horizontal position and thin copper leads are attached to it at points C and D near to its ends. The wire is connected in series with an adjustable resistance, a battery, a key K, and a coil S. A small current, not sufficient to raise the temperature of the wire appreciably above that of the room, is sent through it. The potential differences between C and D and across S are measured with the aid of the millivoltmeter MV. Mercury cups, *a* . . . *h*, in a block of paraffin wax enable the connections to be made easily. From these observations the

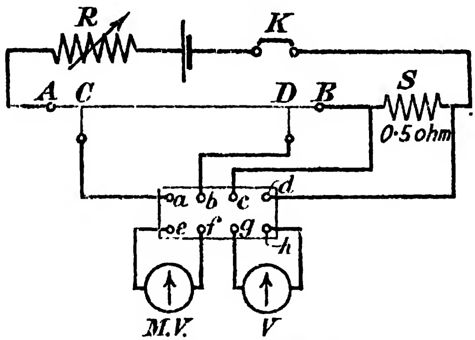


FIG. 48·6.—Thermal Emissivity of a Wire.

current in the circuit and the resistance of the portion CD of the wire are deduced. A current of about 1·5 amperes is then passed through the wire and the potential differences between the same points determined with the aid of the voltmeter V. The current and resistance are again calculated. From a knowledge of the coefficient of increase of resistance with temperature for nickel [0·0052 deg.$^{-1}$ C.] the temperature of the wire may be found. This temperature has become steady since the rate at which energy is dissipated in the wire is equal to the rate at which heat is being lost from its surface. The thermal emissivity is calculated as follows.

Example.—

(i) P.D. across 0·5 ohm $= 47{\cdot}4 \times 10^{-3}$ volt.
P.D. „ CD $= 72{\cdot}5 \times 10^{-3}$ volt.
$\therefore$ resistance of CD $= 0{\cdot}765$ ohm at 19·5° C. ($t_1° \simeq \phi_1$)

(ii) P.D. across 0·5 ohm $= 0{\cdot}75$ volt. $\therefore$ current $= 1{\cdot}5$ amp.
P.D. „ CD $= 1{\cdot}75$ volt.
$\therefore$ resistance of CD when heated $= 1{\cdot}17$ ohm $= R_t$.

Since $R_t = R_0(1 + \kappa\phi)$ and $R\phi_1 = R_0(1 + \kappa\phi_1)$, we have

$$\frac{1{\cdot}17}{0{\cdot}765} = \frac{1 + \kappa\phi}{1 + \kappa\phi_1} \qquad \therefore \phi = 131° \text{ C.}$$

Length of wire = 33·8 cm. Diameter = 0·023 cm.

$$\text{Emissivity} = \frac{\text{Amount of heat (cal.) emitted per sec.}}{\text{Area of surface} \times \text{temp. diff.}}$$

$$= \frac{1{\cdot}5^2 \times 1{\cdot}17}{4{\cdot}2 \times \pi \times 0{\cdot}023 \times 33{\cdot}8 \times 111{\cdot}5}$$

$$= 2{\cdot}3 \times 10^{-3} \text{ cal. sec.}^{-1} \text{ cm.}^{-2} \text{ deg.}^{-1} \text{ C.}$$

THERMOELECTRICITY

The Seebeck Effect.—In 1821 SEEBECK discovered that an electric current could be produced by thermal means alone. He showed that a current flows in a circuit consisting of two wires of different materials as long as there is a difference in temperature between the two junctions. We must note, however, that the current ceases as soon as this temperature difference becomes zero. Such currents are termed thermoelectric currents and the electromotive force producing them is known as a thermoelectric force.

One of the first pieces of apparatus whereby Seebeck discovered how to produce thermoelectric currents is shown in Fig. 48·7 (*a*). The ends of a copper bar were bent as indicated, and soldered to a

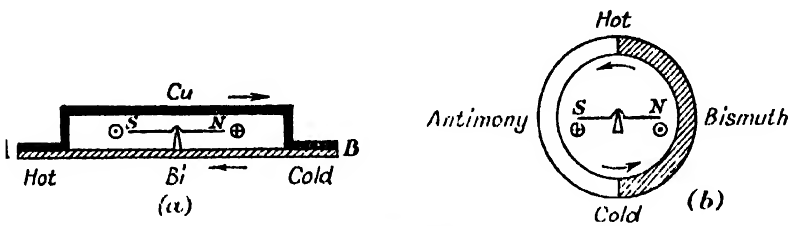


FIG. 48·7.—Seebeck Effect.

bar of bismuth. The plane of this combination of metals was placed in the magnetic meridian, and a small compass needle was supported near to the middle of the circuit. When the junction A was heated the pole N of the magnet was deflected towards the west (this movement is indicated in the diagram by the ⊕ sign), i.e. an electric current had been produced and it flowed from the copper to the bismuth across the cold junction. If the circuit consists of rods of antimony and bismuth, bent for convenience to form a circle, the direction of the current is from the antimony to the bismuth through the cold junction [cf. Fig. 48·7 (*b*)]. The direction of the current may be found by placing the plane of the coil in the meridian and observing the motion of a magnetic needle at its centre. In each instance the energy necessary to maintain the current is derived from the surroundings, i.e., on the whole, there is an absorption of heat from them.

Experiment.—AHB is a rod of copper 1 cm. in diameter and bent as indicated in Fig. 48·8 (*a*). It is short-circuited by C, a thick piece of constantan [Cu 60 per cent., Ni 40 per cent.]. A large iron block is cut in halves, grooved, and fitted round the rod as shown in Fig. 48·8 (*b*). AB is insulated from the iron blocks by paper. When the end A of the rod is heated whilst B is kept in ice it is almost impossible to separate the two pieces of iron. This is because the iron has

become magnetized by the large current in the circuit HKL. This current is produced by the small thermoelectric force in the same circuit which appears when the two junctions K and L are at

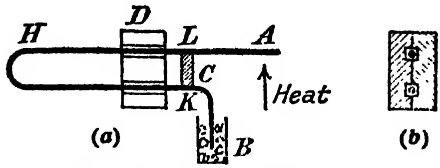


FIG. 48·8.—Seebeck Effect.

different temperatures. The current is large since the resistance of the circuit is very small if the junctions K and L have been well made—they should be silver-soldered.

The Peltier Effect.—An effect which is the converse of that just described was discovered by PELTIER in 1834. He noticed that when a current passed across the junction between two dissimilar metals there was either an evolution or an absorption of heat, i.e. the junction was either heated or cooled. For any two metals the condition whether the junction shall be heated or cooled is determined by the direction of the current. This effect is entirely apart from the Joule effect which is irreversible, i.e. it does not depend on the direction of the current in the conductor, whereas the Peltier effect is reversible. In the following experiment the former effect is made negligibly small by using thick rods. B, Fig. 48·9, is a rod of bismuth soldered at each end to a rod of antimony, A. If

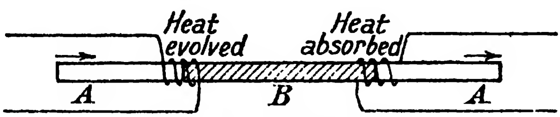


FIG. 48·9.—Peltier Effect.

a current is passed in the direction indicated there is an evolution of heat where the current passes from the antimony to the bismuth. At the other junction heat is absorbed. The effects are reversed when the current is reversed. These effects are clearly shown if coils of thin copper wire are wound round the junctions of the metals. Copper is chosen because its coefficient of increase of resistance with temperature is large. The resistances of these coils are measured when a current is passed along the rods. An increase in resistance of the coil at the junction where the current passes from antimony to bismuth shows that heat is developed at this junction. The other junction is cooled for the resistance of the coil round it decreases. When the direction of the current is reversed, contrary effects are obtained, indicating thereby that the effect under investigation is reversible.

The existence of the Peltier effect may also be demonstrated in the following way. Let T, Fig. 48·10, be a thermopile [cf. p. 325), connected through a key, K_1, to a sensitive galvanometer, G. Further, let us suppose that when heat radiation falls upon the right-hand face of the thermopile the current in the circuit marked (1) is clockwise. Let B be an accumulator, in series with an ammeter A, adjustable resistance R and key K_2, so that a current of about 0·1 ampere may be passed for a period of about 10 seconds through the thermopile T [circuit (2)]. As arranged

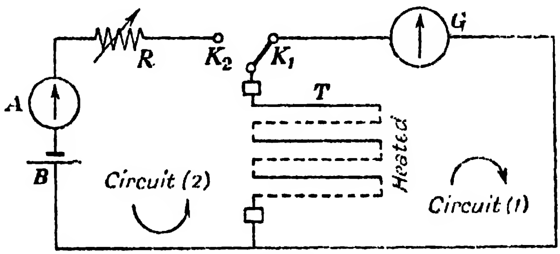


FIG. 48·10.

in Fig. 48·10 the current from B passes through T in the same direction as did the thermoelectric current produced when heat rays fell on the right-hand face of T. When K_2 is opened and K_1 closed immediately afterwards it will be found that G is deflected: the direction of this deflexion will be such as to indicate the presence of an anticlockwise current in circuit (1). Thus the junctions of the thermopile on its R.H.S. must have been cooled with respect to those on the L.H.S. If the current from B through T is reversed, *the current through G when K_2 is closed will also be reversed.*

Both the Seebeck and the Peltier effects may be explained if we assume that there is an electromotive force at the junction of two metals, acting in the case of an antimony-bismuth junction from the bismuth to the antimony, i.e. bismuth is electropositive with respect to antimony. If a circuit consists of two metals only and the temperature is everywhere the same, then the electromotive force at one junction is equal and opposite to that at the other and the total electromotive force in the circuit is zero. If, however, the temperatures of the junctions are different, the opposing electromotive forces are not equal and the difference between them causes a current to flow.

On the other hand, when a current is passed from antimony to bismuth work is done in overcoming the electromotive force at the junction: this appears as heat and the junction is heated. If the current is reversed the junction is cooled. Hence, on these assumptions, the existence of the Peltier effect finds a ready explanation.

In a general way the existence of an e.m.f. at the junction of two metals can be accounted for from our knowledge of the structure of conducting substances. Modern theory suggests that all such bodies contain large numbers of free electrons, i.e. electrons free to move in the interstices between the constituent particles of the conductor. These electrons behave very much like the molecules of a gas so that they are often referred to as an 'electronic gas.' The density of this gas at any fixed temperature of the metal is assumed to vary in different metals, so that when two metals are placed in contact the electrons diffuse from one to the other. This diffusion establishes an electromotive force at the junction which increases in value until it is sufficient to prevent a differential diffusion of the electrons from the one metal to the other. The equilibrium condition finally established is an example of a 'dynamic' equilibrium as distinct from a 'static' equilibrium, for there is no reason to suppose that when equilibrium has been attained the motion of the electrons ceases.

Measurement of the Peltier Coefficient for Two Metals.—*The energy absorbed when unit quantity of electricity passes across the junction of two metals is called the Peltier coefficient for those metals* (Maxwell).—If the energy is measured in joules, the current in amperes, and the time in seconds, the Peltier coefficient will be given in volts. Generally, one of the metals is lead, and if energy is absorbed when the current flows from the lead to the other metal across the junction, the Peltier coefficient is considered positive.

To measure the Peltier coefficient—for a lead-constantan junction, for example—the junction is placed in the inner tube of a Bunsen ice-calorimeter and covered with a small quantity of oil to ensure good thermal contact with the calorimeter. Let a current of I amperes be passed through the junction for t seconds. If R is the effective resistance of the couple, then

$$\text{Heat evolved} = \frac{\mathrm{I}^2\mathrm{R}t + \Pi\mathrm{I}t}{\mathrm{J}} = kx_1,$$

where Π is the Peltier coefficient, J is the mechanical equivalent of heat, viz. 4·18 joules cal.$^{-1}$, k is the constant of the calorimeter determined by an electrical method [cf. p. 232] and x_1 is the distance through which the mercury recedes.

Now let the current be reversed for t seconds. Then

$$\frac{\mathrm{I}^2\mathrm{R}t - \Pi\mathrm{I}t}{\mathrm{J}} = kx_2,$$

where x_2 is the distance the mercury recedes in this instance. Hence,

$$\Pi = \frac{k(x_1 - x_2)\mathrm{J}}{2.\mathrm{I}t} \text{ volt.}$$

[NOTE.—Although the joule heating could be made small and perhaps negligible by using thick leads, this would be a very undesirable procedure in practice since the flow of heat by conduction along the leads to the interior would be augmented and would be very troublesome.]

The Thomson Effect.—From theoretical considerations KELVIN, when he was SIR WM. THOMSON, proved that if the only seat of potential difference in a thermocouple was at the junctions the total e.m.f. in the circuit should be proportional to the temperature difference between the junctions. Experiment shows that this is not even approximately true so that Kelvin assumed that in any homogeneous wire there must be an e.m.f. whenever there is a temperature gradient in the wire. To test the validity of this conclusion, Kelvin sent a heavy current through a homogeneous bar. The ends of this bar were kept at a constant temperature, but the central portion was heated. Then, with the aid of a sensitive differential thermometer, he showed that the amounts of heat generated in the two halves of the bar were unequal. He also showed that the effect was reversible.

Experiment.—A long thin U-shaped piece of iron wire is supported so that the bend dips into a considerable amount of mercury and a current of such strength that the wire is just visible in a darkened room is sent down one limb and up the other. The wire is cooled by the mercury so that there is a temperature gradient in the wire. The two portions of the wire glow unequally, showing that there are opposite Thomson effects in the two limbs.

Electron Theory and the Thomson Effect.—Since the electron density is greater at low temperatures than at high, it follows that we may expect an electromotive force whenever a temperature gradient exists in a conductor. Thus the simple electron theory accounts for the existence of the Thomson effect. Unfortunately, however, this theory would indicate that the e.m.f. is always directed from the region of high temperature to that where it is low, i.e. electrons move from the cold to the hot region, whereas experiment shows that this direction depends upon the material investigated.

Experimental Determination of Thermal e.m.fs.—(*a*) ***Approximate method.*** Let us suppose that we have to investigate the manner in which the e.m.f. of a copper-constantan * thermocouple varies with temperature, one junction being maintained at the temperature of melting ice. A constantan wire is hard soldered to two pieces of copper wire to form two copper-constantan junctions as shown in Fig. 48·11 (*a*). The copper wires are then connected to the terminals of a high-resistance

* Constantan: 60 Cu, 40 Ni. This alloy is also known as eureka.

galvanometer G. It will be remembered that the coil of wire inside a galvanometer is made of copper: hence there are still only two different metals in the circuit. The thermocouples are placed in two different test-tubes surrounded by melting ice and by water respectively. One of the junctions is maintained at the temperature of melting ice throughout the experiment, while the temperature of the other may be altered by heating the water. When the temperatures of the two junctions are different the electromotive force which is developed in the circuit causes a current to pass through the galvanometer. Since G has a high resistance its deflexion may be taken as proportional to the e.m.f. in the circuit. If absolute values of the e.m.f. are required the volt sensitivity of the galvanometer must be known. Usually the current sensitivity, i.e. the current required to produce a given scale deflexion (1 mm. when the scale is at a distance of 1 metre for galvanometers used in conjunction with a lamp and scale), is stated on the instrument, but if the resistance of the galvanometer is known the volt sensitivity is easily deduced.

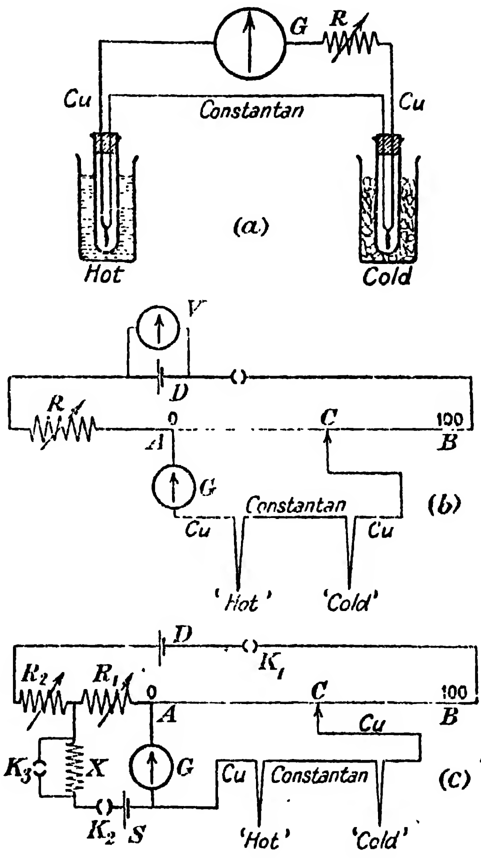


FIG. 48·11.—Experimental Determination of Thermal e.m.fs.

(*b*) ***Simple potentiometer method.***—Since the thermal e.m.f. which it is proposed to measure is small, a wire with a small potential drop per unit length along it is required. Suppose that the potentiometer wire AB, Fig. 48·11 (*b*), has a resistance r ohms and it is placed in series with a resistance R ohms and a battery D of

e.m.f. E. The potential drop across the wire is equal to the current along the wire multiplied by the resistance of the wire, viz.

$$\frac{Er}{R+r}$$

If E is of the order 2 volts, r of the order 10 ohms, and R of the order 2,000 ohms, the potential drop across the potentiometer wire will be 10^{-2} volt, or 10^{-4} volt.cm.$^{-1}$ if the wire is one metre long. One end of the thermocouple is connected to A through a galvanometer G. The other end is connected to a jockey C which slides along the wire AB. The position of C is adjusted until the galvanometer deflexion is zero. The e.m.f. of the thermocouple is then equal to the potential drop across the portion AC of the potentiometer wire, and is given by the expression

$$\frac{Er}{R+r}\cdot\frac{AC}{AB}$$

An objection to this method is that the battery D supplies an electric current and therefore its e.m.f. on open circuit is not the e.m.f. available for sending the current through the circuit. The correction is small, however, for the current supplied by the battery is not large. This available e.m.f. may be measured by placing a voltmeter, V, in parallel with D. Another objection, and one which is more serious, is that there is no means of maintaining a constant current in the circuit. This difficulty is overcome by proceeding as in (*c*).

(*c*) ***Using a standard cell.***—The resistance box R is replaced by two boxes, R_1 and R_2, Fig. 48·11 (*c*), the sum of their resistances being of the same order as that of R, so that the fall of potential along the wire shall still be comparable with that of the thermocouple. S is a standard cell placed in series with a high-resistance X(10,000 ohms) to prevent large currents from being taken from the cell. These are arranged as shown.

With the keys K_1 and K_2 closed, the values of R_1 and R_2 are adjusted so that there is no deflexion of G. The potential drop across R_1 is then equal to the e.m.f. of S on open circuit since this cell is supplying no current. If necessary, when an approximately correct balance has been obtained, the key K_3 may be closed to short circuit X; this permits the galvanometer to be used at its maximum sensitivity. K_2 and K_3 are then opened.

The e.m.f. of the thermocouple is determined by finding the point C on the wire AB corresponding to no deflexion of the galvanometer. If E is the e.m.f. of the standard cell, the fall of potential across R_1 is E, so that the current in the main circuit is E/R_1. The drop in potential across the potentiometer wire is therefore

$$r\left(\frac{E}{R_1}\right),$$

so that the potential difference between A and C is therefore

$$r\cdot\frac{E}{R_1}\cdot\frac{AC}{AB}.$$

This is the e.m.f. of the thermocouple.

[N.B.—It is not necessary to know R_2 and when once the potential drop across R_1 has been made equal to the e.m.f. of S on open circuit, the constancy of the current along the potentiometer wire may be tested and maintained by keeping R_1 constant and adjusting R_2 so that the deflexion of G is zero when K_2 is closed.]

As before one junction of the thermocouple is placed in a test-tube and surrounded by melting ice. The other is heated to different known temperatures and in each instance the position of the sliding contact C when the galvanometer deflexion reading is zero noted. The thermal e.m.f. in the circuit is in each instance equal to the potential drop along the corresponding portion of the wire AC.

The Laws of Thermoelectricity.—Let A and B be two metals forming a thermocouple; θ_1 and θ_2 the temperatures of the cold and hot junctions respectively. The e.m.f. in the circuit will be denoted by $\Big[E\Big]_{\theta_1}^{\theta_2}$ and considered positive when the current flows from A to B across the hot junction. The laws of the thermoelectric circuit are :—

(i) The law of intermediate temperatures: For a given pair of metals the e.m.f. when the junctions are at θ_1 *and* θ_3 *is equal to the sum of the e.m.fs. when the junctions are at* θ_1 *and* θ_2 *and then at* θ_2 *and* θ_3,

i.e.
$$\Big[E\Big]_{\theta_1}^{\theta_3} = \Big[E\Big]_{\theta_1}^{\theta_2} + \Big[E\Big]_{\theta_2}^{\theta_3}$$

(ii) The law of intermediate metals: When the temperatures of the junctions are fixed, the e.m.f. for two metals A and C [denoted by E_{ac}*] is equal to the sum of the e.m.fs. for the thermocouples A, B and B, C, i.e.*

$$E_{ac} = E_{ab} + E_{bc}.$$

The first law may be verified experimentally: the method is so simple that no description is necessary. It is important practically, for if observations are made with a given thermocouple for various temperatures of the hot junction, the cold one being kept at 0° C., then the e.m.f. appropriate to other temperatures of the junctions may be obtained by subtraction, for

$$\Big[E\Big]_{\theta_2}^{\theta_3} = \Big[E\Big]_{\theta_1}^{\theta_3} - \Big[E\Big]_{\theta_1}^{\theta_2}.$$

The validity of the second law may be tested experimentally by the following null method. Metals A, B and C, where C is copper, are arranged to form a 'star-pattern' as in Fig. 48·12. Copper is selected as one of the metals so that a moving-magnet galvanometer (all coils made of copper) may be inserted to detect any current without introducing a different metal. If the 'outer' or hot junctions are all raised to the same temperature and the 'inner' or cold junctions all kept in melting ice, no current flows. Since,

however, the circuit is equivalent to three thermocouples [A, C], [B, A], [C, B] in series, we have

$$E_{ac} + E_{ba} + E_{cb} = 0,$$

i.e.

$$E_{ac} = E_{ab} + E_{bc}.$$

The importance of the law of intermediate metals is that when a thermocouple is constructed it is immaterial whatever alloy is used as a solder to join together the dissimilar metals. [In some instances the metals are twisted together, but this practice sometimes leads to trouble if oxidation occurs, for the resistance of the oxide may be very high.] Further, if the third metal forms the leads to a galvanometer no effect whatsoever is produced provided that the junctions of these leads with the other two metals are at a constant temperature—generally 0° C. in practice.

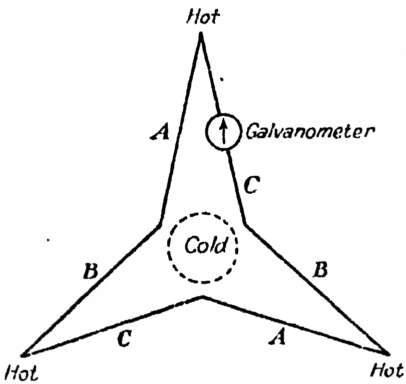


FIG. 48·12.

The e.m.f. of a Thermoelectric Couple.—Experiment shows that for a wide range of temperatures the e.m.f. of a given thermocouple, when one junction is at the temperature of melting ice, is given by the equation

$$E = \alpha\theta + \tfrac{1}{2}\beta\theta^2,$$

where $(\theta + 273)$ is the temperature on the gas-scale, and α and β are constants for a given pair of metals.

To determine the constants α and β for, say, an iron-constantan thermocouple these metals must be joined together with 'easy-flow' solder, and then the free end of each metal must be soldered to copper. The copper-iron and copper-constantan junctions must then be maintained at the same constant temperature, (0° C.), while the iron-constantan junction is placed in baths at known temperatures—say steam, under a known pressure in the region of 76 cm. of mercury, tin at its melting-point, etc. The e.m.f. in each instance must be determined as on p. 928. Then if $\frac{E}{\theta}$ is plotted against θ, a straight line will be obtained, for the above equation becomes

$$\frac{E}{\theta} = \alpha + \tfrac{1}{2}\beta\theta,$$

which is a straight line—slope $\frac{1}{2}\beta$, and intercept on the $\frac{E}{\theta}$ axis is α.

[N.B.—If a mercury thermometer is used in this experiment a correction table must be used to convert its readings to those on the gas-scale].

Thermoelectric Curves.—Let us consider a copper-iron thermocouple one of whose junctions is maintained at 0° C. while the other is raised in turn to a series of known temperatures. Suppose the e.m.f. is measured in each instance and a graph drawn showing the relation between the e.m.f. in the circuit and the temperature of the hot junction. The curve is known as a ***thermoelectric curve***—see Fig. 48·13. In practically all instances it is

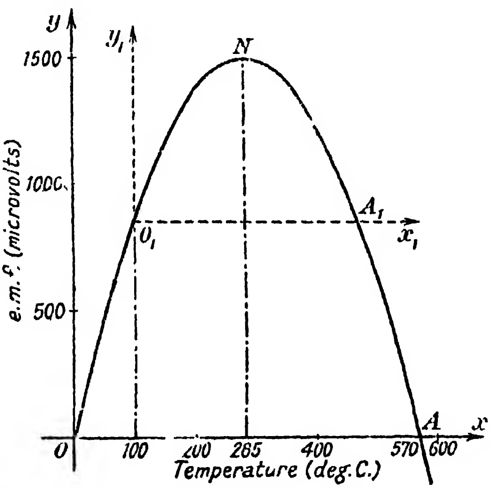


FIG. 48·13.—Thermoelectric Curve for a Copper-Iron Thermocouple.
$E_{Cu\,Fe} = [10{\cdot}34\theta - 0{\cdot}0183\theta^2]\ \mu V.$

a parabola with its axis vertical [in some instances only a portion of the curve is obtainable—one of the wires may melt].

If the temperature of the cold junction is raised to 100° C. (say) the corresponding thermoelectric curve is obtained by transferring the axes of co-ordinates to O_1x_1, O_1y_1 as shown, where the abscissa of O_1 is 100° C.

Thermoelectric Inversion.—A study of the thermoelectric curve for a copper-iron thermocouple shows that at a certain temperature of the hot junction the thermoelectric e.m.f. in the circuit is a maximum. This is termed the ***neutral temperature.*** As the temperature is raised the e.m.f. decreases, becomes zero, and then reverses its sign. The temperature corresponding to the

neutral point is the mean of the temperatures of the cold junction and the temperature at which the above reversal of sign begins.

Experiment.—The above facts are very easily obtained by using a copper-iron thermocouple in series with a high resistance galvanometer. The galvanometer deflexions are proportional to the thermal e.m.f. in the circuit.

Thermocouples used as Thermometers.—Since the e.m.f. of a thermocouple, with one junction at 0° C., depends upon the temperature of the hot junction, it is clear that when the corresponding thermoelectric curve has been obtained that curve may be used to determine an unknown temperature within the range of temperature covered by the curve. For such a purpose a copper-iron couple would not be chosen on account of the ambiguity which arises when the temperature is in the neighbourhood of the neutral point. For temperatures up to about 500° C. (gas scale) thermocouples made from copper-constantan, and iron-constantan, are satisfactory: the e.m.f. is large so that the temperature may be estimated accurately. Other pairs of metals must be used at higher temperatures: these will be discussed later.

Meanwhile it must be noted that any pair of metals may be used to construct a centigrade scale of temperature. The metals selected for such a purpose may be constantan-copper. Let us suppose that the e.m.f. is measured by the first experimental arrangement shown in Fig. 48·11 (*a*), although (*b*) and (*c*) may be used. The junctions of the thermocouple are placed in melting ice and steam produced under a pressure in the neighbourhood of 76 cm. of mercury, respectively. The resistance R is adjusted until the deflexion on the galvanometer scale is about 200 mm.: it is advantageous to make it equal to $2s$, where s is the steam temperature on a centigrade scale as calculated from the barometric height. The galvanometer deflexion is then such that 2 mm. corresponds to one degree on a centigrade scale of temperature. It must be remembered, however [cf. p. 152], that temperatures on such a scale will not in general be identical with those on any other centigrade scale.

Experiment—Set up a thermocouple as above and use it to construct a centigrade scale of temperature; then place one junction in molten wax, observe how the deflexion varies as the wax cools, and from a suitable graph obtain a value for the melting-point of wax on the scale of temperature you have constructed.

Some Applications of Thermoelectricity.—Thermocouples are widely used in industry, since they provide a ready and sufficiently accurate means of measuring temperatures over a wide range. For temperatures below 1200° C. base metals may be used. The construction of a nickel-nichrome thermocouple is indicated in Fig. 48·14. AC and BC are the two wires welded together at C to form the 'hot

junction.' The whole is protected by an iron sheath, the wires being insulated from it and one another by fireclay insulators. The wires

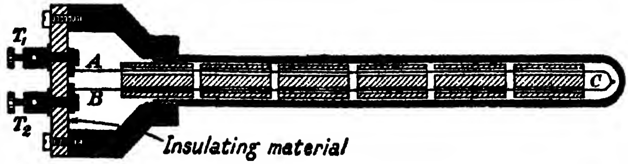


FIG. 48·14.—A 'Base-Metal' Thermoelectric Pyrometer for use to 1200° C.

are connected to terminals T_1 and T_2 fixed in the head of the pyrometer —made from insulating material.

For use at higher temperatures the wires are made from metals of the platinum group, since at these temperatures it is necessary that the metals should be highly infusible and not affected by air. The wires should be very thoroughly annealed before use, and, after prolonged use, re-annealed at a higher temperature than the maximum at which they have been used since platinum readily absorbs gases.

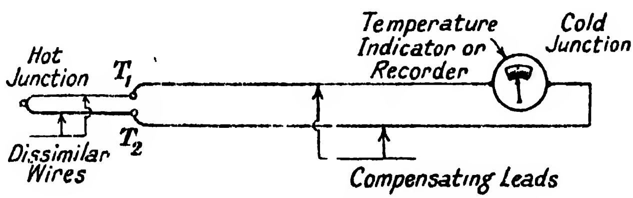


FIG. 48·15.—Thermoelectric Pyrometer.

The sheath and insulators are made of silica. For temperatures from 1400° C. to 1700° C. a couple made from molybdenum and tungsten wires is suitable. At these temperatures, however, the wires are brittle and must therefore be protected from shock; moreover, they rapidly oxidize unless used in an atmosphere of 'cracked' ammonia, i.e. a mixture of nitrogen and hydrogen produced by passing ammonia through a tube heated to 1000° C. Such couples are used to determine the temperature of molten iron and steel.

When used with a temperature indicator, the arrangement is as in Fig. 48·15. 'The cold end' is brought from the head of the thermometer to the indicator, where the temperature is reasonably constant. The leads are made of copper. Since the e.m.f. of a thermocouple depends on the temperature of the 'cold junction,' the needle of the indicator [really a millivolt-meter] is adjusted to 'room-temperature' before being connected to the pyrometer. For accurate work T_1 and T_2 must be kept at the temperature of melting ice.

The use of thermocouples in thermopiles has already been described [cf. p. 325]. Thermopiles, it must be noted, are only used to detect heat radiations: they are not used to measure temperatures.

The Radiomicrometer.—A very sensitive instrument for detecting feeble thermal radiations is known as a radio-micrometer—it was invented by D'ARSONVAL and by BOYS, and is shown dia-

grammatically in Fig. 48·16. CD is a loop of copper or silver wire [No. 36 S.W.G., diameter 0·2 mm.]. The circuit is closed at its lower end by an antimony-bismuth junction, A, B. A piece of the same wire, about 5 mm. long, is soldered to the loop at its upper point, and attached by shellac varnish to a glass capillary tube G. M is a plane mirror, edge about 3 mm. in length. It is a piece of cover glass, selected by optical trial for planeness, and silvered at the back. The loop is suspended by a fine quartz fibre [0·004 mm. in diameter], between the poles of a strong magnet. Since antimony and bismuth are very diamagnetic, they must be screened from the influence of the magnet by surrounding them with a large block of soft iron [not shown in the diagram]. A hole drilled in the iron block permits radiation to fall on a very small piece of blackened copper foil attached to the junction between the antimony and the bismuth. The sensitivity of this instrument is such that when the temperature of the copper disc is raised only by a few millionths of a degree the current in the loop is sufficient for it to be deflected.

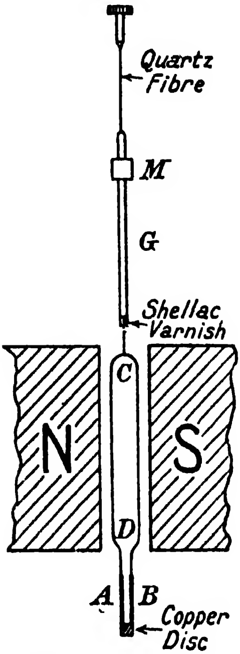


FIG. 48·16.—Boys' Radiomicrometer.

M enables these deflexions to be measured by a lamp-and-scale method, and it is placed about 3 cm. from the lower end of G, so that the heat falling on M from the lamp shall not be troublesome. The wire carrying the loop is torsionally infinitely rigid compared with the quartz fibre, so that any deflexion of the loop is truly measured by the deflexion of M. With such a sensitive instrument great precautions must be taken to screen it from thermal changes—it is enclosed in a wooden box—and one also notices that there are no outside leads which might cause an induced current in the circuit should they 'cut' the earth's magnetic field [cf. p. 936 et seq.]. In fact, the current is detected without the aid of any additional galvanometer! Moreover, with this instrument Boys experienced no trouble due to variations in the external magnetic field, as did Langley, and others, who used bolometers in conjunction with moving magnet galvanometers.

The sensitivity of the above radiomicrometer is such that the radiant energy from a candle flame two miles away may be detected.

EXAMPLES XLVIII

1.—Two wires of 3 and 4 ohms resistance respectively are joined to a battery of negligible resistance, first in parallel, and then in series. Calculate the ratio of the heats developed in the two systems.

2.—A current of 4 amperes flows through a resistance of 6 ohms for 3 mins. If the heat developed is sufficient to raise the temperature of 600 gms. of liquid 7° C., calculate the specific heat of the liquid.

3.—The poles of a battery are connected in turn to two wires of 5·2 and 4·3 ohms resistance respectively. The heats developed in the two wires are equal. Calculate the resistance of the battery.

4.—How would you show that the rate at which heat is developed in a conductor carrying a current is proportional to the square of the current?

5.—How would you demonstrate the Seebeck and Peltier thermoelectric effects? Describe how you would use a thermocouple to measure temperatures up to 1000° C.

6.—Compare the heat generated in each of the four arms of a balanced Wheatstone bridge, if the resistances of those arms are 100, 10, 550, and 55 ohms respectively.

7.—Derive an expression for the rate at which electrical energy is dissipated in a coil of wire of resistance R ohms carryng a steady current of I amperes.

A battery, of e.m.f. 10 volt. and internal resistance 2 ohm., is connected in series with an adjustable resistance R. Plot a graph showing how the rate of dissipation of energy in R varies with the value of R. What information is obtained from such a graph?

8.—Indicate by means of diagrams the modifications and additions required to use a simple potentiometer (*a*) to measure a current of about (i) 0·1 amp., (ii) 2 amp., (iii) 10 amp.; (*b*) to measure the e.m.f. of a thermocouple.

Give reasons why it is permissible to use a potentiometer to measure potential differences in an experiment to verify Ohm's law for a metallic conductor but fallacious to use it to measure currents in the same experiment.

9.—A tangent galvanometer and a coil of insulated resistance wire immersed in a calorimeter containing water are joined in series and a steady current, which gives a deflexion of $\tan^{-1} 0{\cdot}4$ on the galvanometer, is passed for 6 min. 58 sec. The temperature of the calorimeter and its contents rises from 15·80° C. (room temperature) to 22·51° C. while the current is passing and begins to fall at a rate 0·15 deg. C. min.^{-1} when the current is switched off. If the galvanometer has 2 turns of wire of 15·7 cm. radius, $H_0 = 0{\cdot}20$ oersted and the calorimeter of thermal capacity 9·5 cal. deg.^{-1} C. contains 90·5 gm. of water, calculate a value for the resistance of the coil. Give a diagram of the electrical circuit and suggest any modifications which would lead to a more accurate result.

CHAPTER XLIX

ELECTROMAGNETIC INDUCTION

Faraday's Discovery of Electromagnetic Induction.—In 1831 FARADAY discovered that induced currents were set up in a closed circuit whenever a current in a neighbouring circuit was made or broken, i.e. when there was a change in the number of lines of magnetic induction threading the closed circuit. For several years previous to this Faraday had failed to detect the presence of these currents, a fact due to the low sensitivity of the galvanometer he used. From his published account of this work it appears that his first successful attempt was carried out on the following lines. About 200 feet of copper wire were coiled round a large block of wood; a second long length of similar wire was then interposed as a spiral between the turns of the first coil, twine serving as an insulator. One spiral was connected to a galvanometer and the other to a battery. When the battery circuit, the so-called ***primary circuit***, was closed, 'there was a sudden and very slight effect (deflexion) at the galvanometer'—i.e. there was an induced current of a transient nature produced in the galvanometer circuit—the so-called ***secondary circuit***. There was also a similar effect, but in a contrary sense when the primary current was broken. Faraday is very careful to emphasize the fact that the current in the secondary circuit is a transient one and that no current exists there when the current in the primary is fully established.

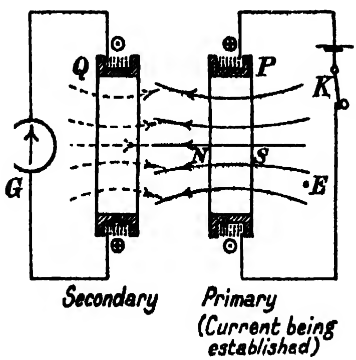


FIG. 49·1.—Faraday's Discovery of Electromagnetic Induction.

The above results may be verified in the laboratory in the following manner. P, Fig. 49·1, is the primary coil connected to a battery and a key K. Q is the secondary coil connected to a ballistic galvanometer G. It will be found that when K is closed

there is an induced current in the secondary circuit: also when the primary current is broken. Both these currents are of short duration but opposite in direction. It will also be noticed that there is no current in the secondary circuit when that in P is fully established. If, however, the current in P is increased there is a transient current in Q—it is in the same direction as that established when the primary is first closed. If the primary current is reduced, an induced current in the opposite direction is temporarily established in Q.

Let us suppose that the current in P is such that to an observer at E it appears to flow in a clockwise direction. Then the lines of magnetic induction are as indicated. Then the current in Q, when that in P is being established, is such that it appears to flow in an anti-clockwise direction to an observer at E—the lines of magnetic induction are shown by the dotted curves. The induced current is such that it tends to maintain constant the number of lines of magnetic induction threading the primary coil—this is a general principle applicable to all induced currents.

Faraday then modified his secondary circuit as follows—the galvanometer was replaced by a small hollow helix of copper wire wound on a glass tube. In this he inserted an unmagnetized steel needle, the primary circuit being open. The primary circuit was then closed, and on removing the steel needle it was found to be magnetized. This was further evidence that a current had been established in the secondary circuit. He varied the experiment by first establishing the primary current, then placing the needle in the helix, and afterwards breaking the primary circuit—the needle was again magnetized but with the direction of the magnetic axis reversed.

Faraday also showed that if the secondary circuit was closed after the current in the primary had been established, or varied in any way, no effects were obtained.

Further experiments on electromagnetic induction were as follows. Several feet of copper wire were stretched on a board in the form of a letter W. A similar wire was then erected on a second board, so that when the two were brought together there would have been contact at all points had not a thick sheet of paper been interposed. One wire was connected to a battery and the other to a galvanometer. On causing one circuit to approach the other a transient current was established in the galvanometer circuit—a transient current in the opposite direction was obtained when the distance between the two circuits was increased.

In later experiments by Faraday a small permanent current was introduced into the galvanometer circuit—the deflexion being about 30°. Transient currents, shown by the temporary excursion of the

needle from the above position, could be established by any of the methods he had previously described, but in all instances the needle returned to its standard position when the change in the primary current was complete. From all these experiments Faraday had not the least doubt concerning the true nature of the effect he had discovered, but states that he was not able to detect the presence of these currents by his tongue, by a heating effect, or by a chemical effect, 'though the contacts with the metallic or other solution were made and broken alternately with those of the battery so that the second effect of induction should not oppose or neutralize the first.' He surmised that failure in this respect was due to the brief duration and the feeble intensity of the induced current.

Further Experiments on Electromagnetic Induction : Faraday 'On the Evolution of Electricity from Magnetism.'—One part of a soft round iron ring—six inches in external diameter—cf. Fig. 49·2—was covered with a helix, P, of copper wire, twine separating the coils in any one layer, and calico separating one layer from the next. A second helix, S, was wound on the other portion of the ring. P was connected to a battery, B, and S to a galvanometer, G. When the current in P was established the galvanometer was immediately affected 'to a degree far beyond what has been described when the helices without an iron core were used, but although the current in the primary was continued, the effect was not permanent, for the needle soon came to rest in its natural position, as if quite indifferent to the attached electromagnet.' When the primary current was broken, the needle was deflected in the opposite direction.

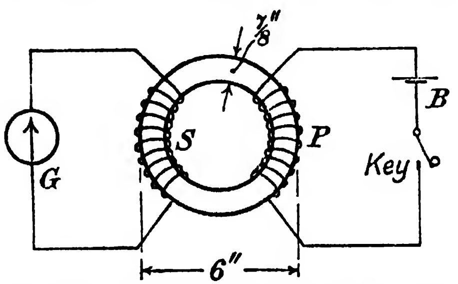


FIG. 49·2.—Apparatus for Producing Induced Currents (after Faraday).

Faraday continues by saying that if matters were arranged so that the direction of the primary current was reversed, the induced currents were contrary in direction to those obtained above, in fact he writes, 'but the deflexion on breaking the battery circuit was always the reverse of that produced by completing it.'

Similar effects were then produced by using ordinary magnets. Among the various experiments carried out by Faraday in this connexion, only the following will be described. A copper helix was wound on a pasteboard cylinder, 8·5 in. long and 0·75 in. in diameter. This coil was connected to a galvanometer. On intro-

ducing a cylindrical magnet into the helix—cf. Fig. 49·3—the galvanometer needle being stationary, the needle was deflected, but having been thus introduced, the needle returned to its zero position. When the magnet was withdrawn the deflexion was in the opposite direction.

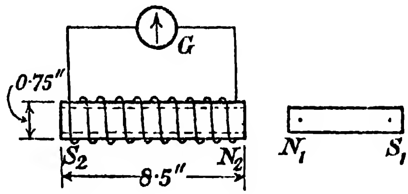


FIG. 49·3.—Faraday's Apparatus for Producing Induced Currents by the Motion of a Bar Magnet near to a Closed Coil

In the above experiment the magnet must not be passed entirely through the helix for a second action then occurs. When the magnet is introduced the galvanometer needle exhibits a certain deflexion, but, being in, a deflexion in a direction contrary to that obtained initially occurs when the magnet is withdrawn, or if it is pushed right through the helix. If the magnet is passed right through in one continuous movement, the needle moves one way, is stopped, and finally moves the other way.[1]

The above experiment may be repeated using the apparatus shown in Fig. 49·4 (*a*). The diagram also shows the direction of

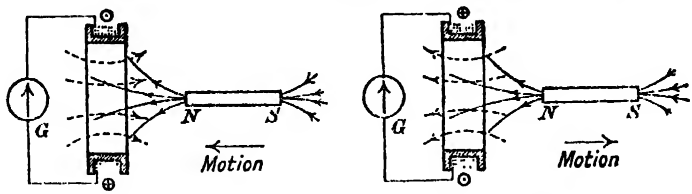


FIG. 49·4.—Induced Currents Produced by the Motion of a Bar Magnet near to a Closed Coil.

the lines of magnetic induction due to the magnet NS and also the lines of magnetic induction due to the induced current when NS approaches the coil. The direction of the induced current is such that the number of lines of induction threading the coil tends to remain constant. Fig. 49·4 (*b*) shows the direction of the induced current and its associated lines of magnetic induction when NS is being drawn away from the coil.

Experiment.—Fig. 49·5 (*a*) represents schematically an apparatus by means of which the production of induced currents is strikingly shown. A is a closely wound coil consisting of about twelve turns of

[1] If a magnet is passed very rapidly through a coil connected to a galvanometer, no deflexion is obtained—the two effects associated with the entrance and exit of the magnet are over before the magnet (or coil) of the galvanometer has had time to move. The two impulses received are equal and opposite, and the galvanometer does not respond.

thick copper wire. The coil is about 1·5 cm. in diameter. Its ends dip into mercury cups, B and C. A second similar coil, L, with its axis at right angles to the plane of the diagram and its ends dipping into the same mercury cups is supported by a glass rod, D, resting on two other rods, F and H. These are normal to D so that the suspended coil is free to oscillate in a plane at right angles to the diagram. To increase the inertia of the moving system a lead weight, K, may be placed on top of the lower coil, a piece of cardboard serving to prevent the coils from being short-circuited. M is a cobalt steel magnet placed as shown. When the coil L is caused to swing there is a change in the number of lines of magnetic induction linked with it, so that an induced current is produced in the coils which form a closed circuit. Since these consist of thick copper wires their resistance is small and the current large, so that a small magnetic needle, N, placed near the upper coil oscillates with a period of swing equal to that of L.

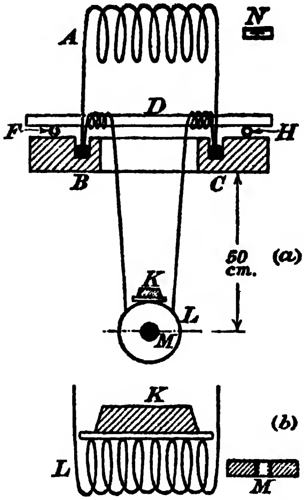


FIG. 49·5.—Experiment on Induced Currents.

Magnetic Flux.—If B_n is the normal component of the magnetic induction at every point of an area A, then B_nA is the ***flux of magnetic induction*** [or ***the magnetic flux***] across that area, and is denoted by Φ. If μ is the permeability of the medium, $B_n = \mu H_n$, where H_n is the normal component of the magnetic intensity. Hence, if the medium is air [strictly, a vacuum], the flux is H_nA, since the permeability of air is unity.

If the magnetic induction is not uniform, the flux is given by $\int B_n . dS$, where the integral extends over the area in question. The unit of magnetic flux is the ***maxwell.***

Magnetic Flux across a Closed Boundary due to an Isolated Pole.—Let a magnetic pole of strength m be situated at O, Fig.

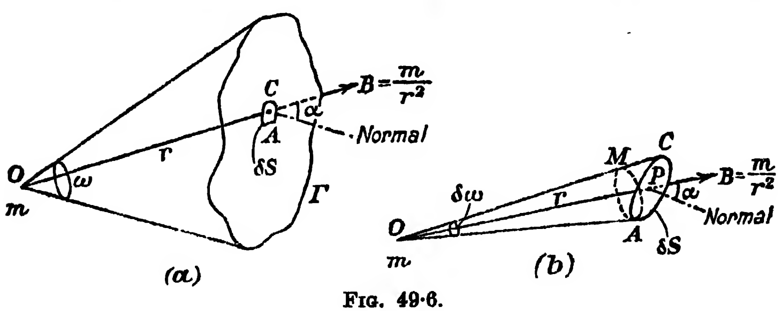


FIG. 49·6.

49·6 (*a*) and suppose that Γ is the closed boundary of a surface across which the flux is to be evaluated. Consider any small element δS of the surface—enlarged in Fig. 49·6 (*b*)—r being the distance of any point in this element from O. Let AM be the projection of AC on a plane through A normal to OP. Let the normal to the element make an angle α with the direction of the magnetic induction B. Then

$$B_n = B\cos\alpha = \frac{m}{r^2}.\cos\alpha.$$

$$\therefore\ B_n.\delta S = \frac{m}{r^2}.\cos\alpha.\delta S = m.\delta\omega, \qquad [\because\ AM = AC.\cos\alpha]$$

where $\delta\omega$ is the solid angle δS subtends at O.

$$\therefore \int B_n.dS = \int m.d\omega = m\omega,$$

where ω is the solid angle which the boundary Γ subtends at O. If the surface surrounds the magnetic pole, m, completely, $\omega = 4\pi$, so that $\Phi = 4\pi m$.

Magnetic Flux and Linkages.—The flux through a boundary has already been defined as $\int B_n.dS$. When the circuit consists of N turns, very close together, the flux through each is the same, but the ***effective flux*** is $N\int B_n.dS$. This is denoted by $N\Phi$, or Ψ, and is called the number of ***linkages*** associated with the circuit; it is expressed in ***maxwell-turns.***

Lenz's Law.—The facts stated previously with regard to the production of induced currents were summarized by LENZ in a law bearing his name. As modified by MAXWELL, it may be stated as follows: ***The e.m.f. induced in a circuit, when a magnet moves in its neighbourhood or the current in an adjacent circuit varies, tends to produce a current, the magnetic field due to which opposes any change in the value of the magnetic flux linked with that circuit.***

Lenz's law may also be expressed as follows: ***When the magnetic flux linked with a closed circuit is changing, there is set up in the circuit a current which tends to keep constant the magnetic flux linked with it.***

Fleming's Right-hand Rule.—When a conductor moves in a magnetic field the directions of the motion, the field, and the induced e.m.f. are given by the following statement due to FLEMING:—***If the thumb and first two fingers of the right hand are spread out so that they point in three directions at***

right angles to one another, the First finger giving the direction of the magnetic Field, the thumb indicating the direction of the Motion of the conductor, then the second finger indicates the direction of the induced e.m.f.—cf. Fig. 49·7 (*a*).

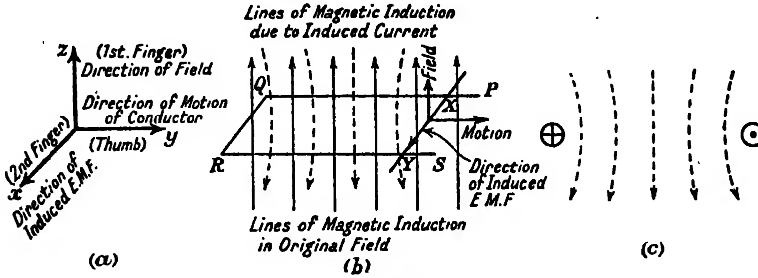


FIG. 49·7.

[Strictly speaking, this rule is only applicable when the magnetic field is normal to the plane in which the displacement occurs. When it is not, the first finger must point in the direction of the component of the field normal to the plane in which the conductor moves.]

Fleming's Right-hand Rule deduced from Lenz's Law.—Let PQRS, Fig. 49·7 (*b*), be a system of long wires, connected as indicated, and lying in a plane normal to the lines of magnetic induction in the medium. Let XY be a conductor bridging the arms of the above system. Suppose XY moves to the right. Then there is a tendency for the flux of magnetic induction through the closed circuit XQRY to increase. The induced current, by Lenz's law, will be such that the magnetic flux due to it tends to prevent the above increase, i.e. the lines of induction [magnetic intensity if the system is in a vacuum] will be downwards [dotted in the diagram]. The current in XY must therefore flow from X to Y—cf. Fig. 49·7 (*c*). This direction coincides with that expressed by Fleming's right-hand rule.

Faraday's Law of Electromagnetic Inductions.—Faraday, by 1831, had established experimentally the conditions under which electromagnetic induction occurs. In 1845 NEUMANN expressed the results of Faraday's work by means of a formula giving the magnitude of the induced e.m.f. In words, it states that the e.m.f. in a circuit is equal to the rate at which the number of linkages, $N\Phi$, associated with that circuit is diminishing. Faraday's law, as given by Neumann, is

$$e = -\frac{d}{dt}(N\Phi) = -\frac{d\Psi}{dt},$$

where e is the induced e.m.f.

Deduction of Faraday's Law from the Principle of the Conservation of Energy.—Helmholtz, and a little later Kelvin, showed that Faraday's law was a direct consequence of the principle of the conservation of energy. Their work appeared about 1850. Let us suppose that a single magnetic pole, of strength m, lies at a point O on the north side of a closed circuit of one turn in which there is a current i (e.m.u)—cf. Fig. 49·8 (a). If ω is the solid angle subtended at O by the periphery of the circuit, then the magnetic potential at O is $i\omega$ [cf. p. 813 et seq.]. This, of course, means that $i\omega$ is the work done in bringing a unit pole from infinity to O. Consequently the work done in bringing up a pole m from infinity to O to $mi\omega$, and this is the potential energy of the system. Now the flux, Φ, through the circuit is $m\omega$ [cf. p. 941].

Suppose that the pole m is moved in time δt to a point such that ω becomes $\omega + \delta\omega$. The work done on the pole is the increase in

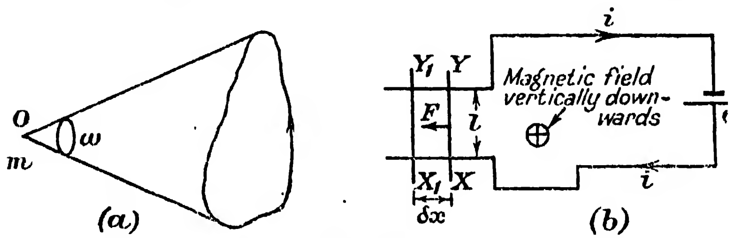


FIG. 49·8.

the potential energy of the system, viz. $mi.\delta\omega$; the rate at which work is done is therefore $mi.\dfrac{d\omega}{dt}$. As the pole moves, let the current in the circuit be maintained constant by increasing the e.m.f. of the generator from E to (E + ε). The additional rate of supply of energy from the generator is εi. By the principle of the conservation of energy we have, therefore,

$$mi.\frac{d\omega}{dt} = \varepsilon i$$

or

$$\varepsilon = m\frac{d\omega}{dt} = \frac{d\Phi}{dt},$$

where $\Phi = m\omega$, the magnetic flux. The electromotive force of induction, e, which is equal and opposite to E, is therefore equal to $-\dfrac{d\Phi}{dt}$. Further, if there are N turns in the circuit, the above e.m.f. will be induced in each of the N turns, so that

$$\text{Induced e.m.f.} = e = -\mathrm{N}\frac{d\Phi}{dt} = -\frac{d\Psi}{dt},$$

where $\Psi = \mathrm{N}\Phi$, the effective flux or number of linkages associated with the circuit.

It will be observed that the induced e.m.f. is independent of i, and therefore has the same value when $i = 0$, i.e. when the initial current is zero.

The above analysis applies when the circuit is fixed and the magnetic field is changing. Now let us consider the instance of a constant field and a moving circuit. To make the analysis as simple as possible we shall assume that the circuit is a plane one and that the magnetic field is normal to this plane at all points. Moreover, the field will be in air, so that $B = H$, the magnetic intensity. Let XY, Fig. 49·8 (b), be a light conductor of length l bridging the arms of a circuit in which there is a cell of e.m.f. e_0 (e.m.u). The mechanical force, F, on the wire is liH acting in the direction shown. Suppose that this force causes the wire to move to X_1Y_1, a distance δx in time δt. The work done by the cell in effecting this displacement is

$$F.\delta x = liH.\delta x.$$

Now in time δt, the energy supplied by the cell is $e_0 i.\delta t$, and if r is the resistance (e.m.u) of the circuit the heat generated therein is $i^2 r.\delta t$ (in work units). By the principle of the conservation of energy

$$e_0 i.\delta t = i^2 r.\delta t + liH.\delta x,$$

so that
$$i = \frac{e_0 - lH.\dfrac{dx}{dt}}{r} = \frac{e_0 - lHv}{r},$$

where v is the velocity with which XY is moving. [If the circuit lies in a medium of permeability μ, H must be replaced by B, the magnetic induction.]

In other words, in addition to the electromotive force of the cell there is an additional e.m.f. equal to $-lHv$, which is minus the rate at which the flux through the circuit is increasing. In this instance flux = number of linkages, so that again we may write

$$e = -\frac{d}{dt}(N\Phi) = -\frac{d\Psi}{dt}.$$

If the moving conductor is not part of a circuit with finite resistance and it moves across a magnetic field as above, there will be a potential difference $-lHv$ between its ends. To establish this, let r be the resistance of the moving conductor and suppose its ends are joined to a high resistance r_1. The current in the circuit will be $-\frac{d}{dt}(N\Phi) \div (r + r_1)$, so that the potential difference across r_1 is

$$-\frac{d}{dt}(N\Phi).\frac{r_1}{r_1 + r} = -\frac{d}{dt}(N\Phi), \quad \text{if } r_1 \to \infty.$$

This establishes the proposition.

Note on Units.—The expression $e = -\frac{d}{dt}(\mathrm{N}\Phi) = -l\mathrm{H}v$ gives the induced e.m.f. in e.m.u. To obtain the value in volts, the expression must be multiplied by 10^{-8} [cf. p. 841].

An inspection of the diagrams shows that the direction of the induced e.m.f. is correctly expressed by Fleming's right-hand rule.

The Quantity of Electricity that flows round a Circuit when the Effective Magnetic Flux linked with a Circuit changes.—It has just been shown that if the effective magnetic flux linked with a circuit is changing at a rate $\frac{d\Psi}{dt}$, there is an induced e.m.f. in the circuit equal to $-\frac{d\Psi}{dt}$. If r is the resistance of the circuit [in e.m.u. of resistance], the induced current will be given by

$$i = \frac{e}{r} = -\frac{1}{r}\frac{d\Psi}{dt},$$
$$= -\frac{1}{r}\frac{d}{dt}(\mathrm{N}\Phi).$$

The quantity of electricity set in motion in time δt is $i.\delta t = \delta q$, say. Hence

$$\delta q = -\frac{1}{r}\frac{d\Psi}{dt}.\delta t = -\frac{\delta\Psi}{r}$$
$$= -\frac{\delta(\mathrm{N}\Phi)}{r}.$$

When Ψ is the total change in the effective flux of magnetic induction linked with a circuit and this change occurs in a time t, we may imagine this interval divided up into a large number of small intervals. If $\delta\Psi_k$ is the change in Ψ in the k-th interval the corresponding quantity of electricity which passes is $q_k = \frac{\delta\Psi_k}{r}$. Hence q, the total charge of electricity which passes in the time t is expressed by

$$q = \sum\frac{\delta\Psi_k}{r} = \frac{\Psi}{r}.$$

Alternatively, the total quantity of electricity set in motion when the number of linkages changes from Ψ_1 to Ψ_2 is given by

$$q = -\int_1^2\frac{d\Psi}{r} = \frac{\Psi_1 - \Psi_2}{r} = \frac{\Psi}{r} \qquad [\text{if } \Psi = \Psi_1 - \Psi_2]$$

[These expressions are only valid if e.m.u. are used. If r is in ohms, it must be remembered that 1 ohm $= 10^9$ e.m.u. of resistance.]

The above expressions for q show that the total quantity of electricity induced is independent of the rate at which the linkages change, but inversely proportional to the resistance of the circuit. Moreover, if this quantity is to be measured with the aid of a ballistic galvanometer the change must be completed before the suspended magnet (or coil) of the galvanometer has moved from its zero position.

The Production of a Continuous Current (d.c.) by means of Electromagnetic Induction.—Faraday placed a copper plate or disc A, Fig. 49·9 (*a*), capable of rotation about a horizontal axis, so that a portion lay between the poles of a strong magnet. The rim of the disc and its axle were well amalgamated so that there was good metallic contact between these parts and copper wires

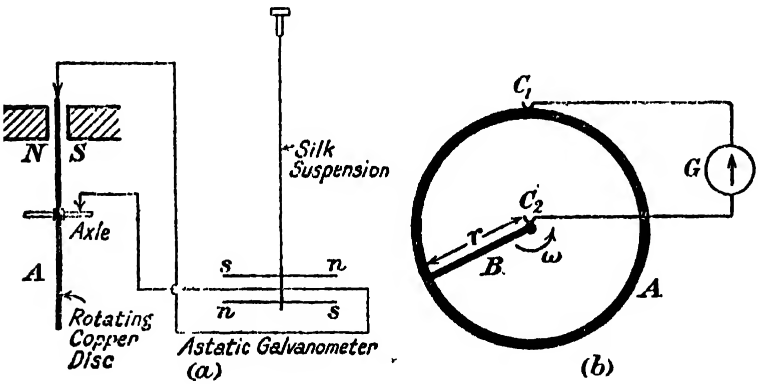


FIG. 49·9.
(*a*) Production of d.c. by means of a Rotating Disc.
(*b*) Disc Rotating at Right Angles to a Magnetic Field.

leading from them to an astatic galvanometer. When the disc was stationary ' the galvanometer exhibited no effect. But the instant the plate moved, the galvanometer was influenced, and by revolving the plate A quickly the needle could be deflected 90° or more. Here, therefore, was demonstrated the production of a permanent current of electricity by ordinary magnets.' When the direction of the disc's rotation was reversed the current was also reversed. This was a very important experiment for Faraday had really made the first ***dynamo.***

In 1832 Faraday modified the above experiment by dispensing with the magnet, NS, and using the earth's magnetic field. The disc was rotated in a plane perpendicular to the direction of the total magnetic intensity and Faraday detected the induced current.

To follow the production of this current more closely, let A,

Fig. 49·9 (*b*), be a circular metal rim of negligible resistance. Suppose that a 'spoke,' B, of length r, revolves with uniform angular velocity, ω, about an axis normal to the plane of the rim and passing through its centre. Let the free end of the spoke be in contact with the rim. Suppose that there is a uniform magnetic field, H, normal to the plane of the rim. Let C_1 and C_2 be two brush contacts connected to a galvanometer G. If the spoke rotates with uniform angular velocity ω, the increase in linkages through the closed circuit comprising the spoke, a portion of the rim, the galvanometer, and the leads to it, is, in time t,

$$H(\tfrac{1}{2}r^2\omega \, . \, t) = \Psi(\text{say}).$$

[H is numerically equal to the magnetic induction, B, if the disc is in air.]

$$\therefore \frac{d\Psi}{dt} = \tfrac{1}{2}Hr^2\omega.$$

If the disc makes N revolutions per second, $\omega = 2\pi N$, and

$$\frac{d\Psi}{dt} = NH(\pi r^2) = NHA,$$

where A is the area of the disc. The magnitude of the induced e.m.f. is therefore expressed by

$$|e| = NHA \quad \text{e.m.u. of potential difference,}$$

$$\text{or} \quad |E| = NHA \times 10^{-8} \text{ volt.}$$

In practice there is no difference between the above and a solid disc revolving with its plane normal to the magnetic field H.

A Quantitative Study of Faraday's Law $e = -\dfrac{d\Psi}{dt}$.— A long solenoid, carrying a steady current, is placed normal to the magnetic meridian. At its centre there is a copper disc C, Fig. 49·10, rotating about a horizontal axis O_1O_2. The plane of C lies

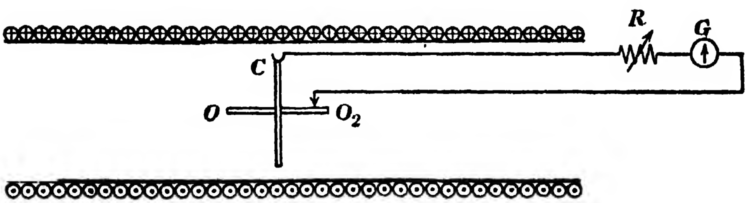


FIG. 49·10.

in the meridian and its area is A. Copper brushes touch the periphery and axle of C and are connected to a known resistor R and microammeter G.

If H is the magnetic field due to the current in the solenoid, A

the area of the disc, n the number of revolutions per second, it has been shown above that

$$|E| = n\text{HA} \times 10^{-8} \text{ volt.}$$

$$\therefore\ I = \frac{n\text{HA} \times 10^{-8}}{R} \text{ amp.} = \frac{n\text{HA} \times 10^{-2}}{R}\ \mu\text{A},$$

where R denotes the total resistance of the circuit. If for different values of n, H, and R, this agrees with the current indicated by G, the expression $|E| = \frac{d\Psi}{dt} \times 10^{-8}$ volt., will have been verified.

Some Calculations Based on Lenz's Law.—(i) A copper disc 20 cm. in diameter rotates about an axis in a plane normal to H_0, the horizontal component of the earth's magnetic field. If the disc makes 5 revolutions per second, calculate the p.d. between the axle and the periphery of the disc.

From the above theory

$$|E| = 5 \times 0{\cdot}18 \times \pi \times 10^{2} \times 10^{-8} \text{ volt}$$
$$= 2{\cdot}8 \times 10^{-6} \text{ volt.}$$

(ii) A vertical copper rod 50 cm. long moves in a plane normal to H and from east to west with a velocity of 100 Km.hr.$^{-1}$ What is the p.d. between the ends of the rod?

An application of the right-hand rule shows that the potential is greatest at the upper end of the rod.

Now in 1 sec. the conductor sweeps out an area

$$\frac{50 \times 10^{5}}{3600} \text{ cm}^{2}.$$

Since $H_0 = 0{\cdot}18$ oersted (this is numerically equal to the magnetic induction if we suppose the motion is in air),

$$\text{e.m.f.} = \frac{50 \times 10^{5}}{3600} \times 0{\cdot}18 \times 10^{-8} = 2{\cdot}5 \times 10^{-6} \text{ volt.}$$

Eddy Currents.—Currents are not only induced in closed wire circuits when the number of lines of magnetic induction threading them varies but also in any conducting material placed in a varying magnetic field. These are termed *eddy* currents. These currents are frequently the source of much trouble in metal apparatus placed in a varying magnetic field. They may cause the metal to become very hot. This may be avoided to a great extent by building up the apparatus from flat metal strips insulated from one another so that the currents are reduced in magnitude.

In recent years advantage has been taken of these eddy currents to melt metals. The specimen is placed in a magnetic field which may pass through from 2000 to 10^7 cycles per second. The field of lower frequency is produced mechanically while the latter is obtained with the aid of a thermionic valve. Not only does the melting take place rapidly, but alloys hitherto unobtainable may be prepared by placing the constituent metals in a high vacuum. Under such conditions the metals do not oxidize and an alloy may be formed.

Experiment (i).—Place an aluminium ring over a solenoid through which an alternating current is passing. If the ring is free to move it is thrown violently off—if it is fixed it becomes considerably heated.

Experiment (ii).—Suspend a copper disc between the poles of an electromagnet. When the magnet is excited the disc may only be moved with difficulty and one experiences the sensation of forcing the disc through a very viscous medium. If the disc moves downwards as indicated in Fig. 49·11, the currents produced have the directions indicated. The magnetic fields associated with these are such that they oppose the cause producing them, i.e. they are such that the electromagnet tends to check the motion of the disc.

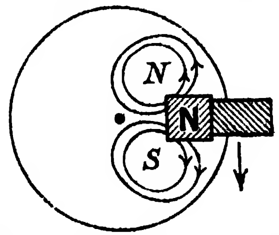


FIG. 49·11.—Eddy Currents produced in a Metal Disc moving in a Magnetic Field.

Experiment (iii).—Allow a magnetized needle to oscillate in turn over a glass sheet and then over a sheet of copper. The oscillations die away more rapidly in the second instance owing to the eddy currents induced in the metal.

Arago's Disc.—The last experiment is due to ARAGO who is also responsible for the following:—A copper disc, Fig. 49·12 (*a*) situated below a magnetized needle NS, was made to rotate rapidly about a vertical axis through its centre. The disc was placed in a box so that the needle was screened from air currents caused by the motion of the disc. The needle was mounted on a pivot fixed to the glass lid of the box. On rotating the copper disc the magnet was deflected from its zero position and tended to move in the

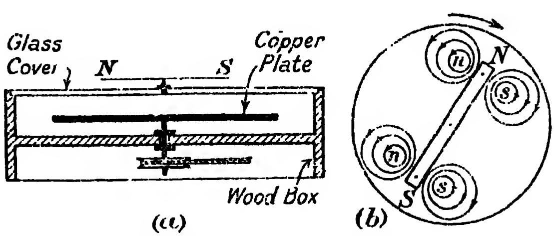


FIG. 49·12.—Arago's Disc.

direction of rotation of the disc: if the speed of the latter were increased sufficiently the needle rotated continuously. The motion of the needle was caused by the eddy currents produced in the disc. These are shown in Fig. 49·12 (*b*). Now the magnet NS was acted upon by a couple due to its presence in the horizontal magnetic field of the earth and by a couple due to the magnetic field caused by the induced currents in the copper plate. For continuous rotation of the needle this latter couple must be greater than the former, i.e. the plate must be given a high angular velocity.

If the angular velocity were below a certain critical value, the needle was only deflected from its standard position and did not rotate continuously.

Arago first carried out this experiment in 1825, i.e. before Faraday had discovered how to induce currents in a circuit. Faraday gave the correct explanation.

As a modification of this experiment, the copper disc may be spun about an axis normal to its plane and passing through its centre, between the poles of the electromagnet: when the magnet is not excited the disc spins easily, but it can only be made to revolve with difficulty when the field is present. This is because the eddy currents in the disc tend to stop its motion and if the disc is made to rotate its temperature increases considerably.

Method Adopted to Diminish Eddy Currents.—Fig. 49·13 (*a*) shows an apparatus devised by WALTENHOFEN. It is essentially a pendulum and a strong electromagnet. The former consists of a copper plate supported so that it may move in a plane between the poles of the magnet. When this is not excited the pendulum swings freely after being displaced. If, however, a field of about 2000 oersted is established between the poles the motion of the plate is very highly damped—in fact it may be dead-beat. Suppose now that a similar plate of copper is cut up into thin strips and that

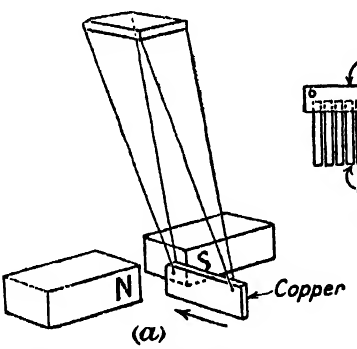


FIG. 49·13.—Waltenhofen's Pendulum.

these, mounted in a wooden frame [cf. Fig. 49·12 (*b*)], are supported as before. The motion of the pendulum is less damped, i.e. the formation of large eddy currents has been prevented.

Similarly, if an iron rod forms the core of a solenoid carrying alternating current, the iron is rapidly heated; when the core consists of sheets of iron, insulated from one another by paper, the heating effect is diminished. [The solenoid should be made from thick copper wire to diminish the Joule effect in it, and the frequency of the current high to increase the magnitude of the eddy currents.]

The Earth Inductor.—When a closed coil is rotated in a magnetic field there is a continuous change in the number of magnetic lines of induction linked with it, so that a current flows in it. This current only lasts whilst the coil is moving and varies from one instant to the next.

Theory.—Let A be the effective area of the rotating coil, i.e. the area of each turn times the number of coils if they are all equal, or the sum of the areas of all the turns if they are unequal, and let this coil make an angle θ with a direction at right angles to that of a uniform field H—cf. Fig. 49·14 (*a*). Then the number of linkages through the coil

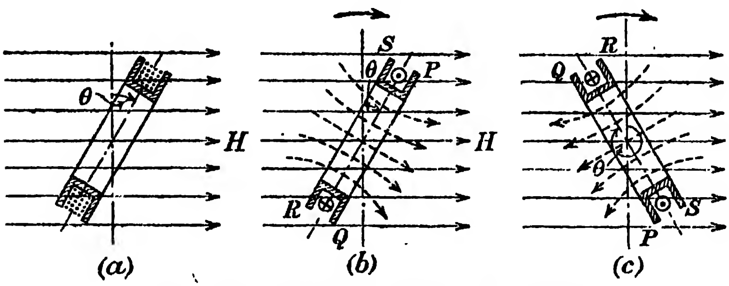


FIG. 49·14.—Principle of an Earth Inductor.

is HA cos $\theta = \Psi$. The instantaneous value of the e.m.f. is therefore $-\frac{d\Psi}{dt} = -\frac{d}{dt}(\text{HA} \cos \theta) = -\text{HA}\frac{d}{dt}(\cos \theta)$. If r is the resistance in e.m.u. of the circuit [the coil and detecting galvanometer, etc.], the instantaneous current is

$$-\frac{1}{r} . \text{HA} . \frac{d}{dt}(\cos \theta).$$

The quantity of electricity flowing in time dt is

$$i . dt = -\frac{\text{HA}}{r} d(\cos \theta).$$

To determine the quantity of electricity, q, passing as the coil is rotated through half a complete turn from a position at right angles to H we must integrate the above expression from $\theta = 0$ to $\theta = \pi$, i.e.

$$q = -2\int_0^{\frac{\pi}{2}} \frac{\text{HA}}{r} . d(\cos \theta) = \frac{2\text{AH}}{r}.$$

If σ is the throw of the ballistic galvanometer when the coil is rotated as above,

$$\frac{2\text{AH}}{r} = \kappa\sigma, \text{ or } \text{H} = \frac{\kappa\sigma r}{2\text{A}}$$

where κ is the reduction factor for the galvanometer.

To determine the direction of the induced current when the rotating coil is in the position shown in Fig. 49·14 (*b*), Lenz's law may be applied. Now the magnetic flux through the coil at this instant is decreasing, so that, by the above law, the induced current

must be such that it tends to increase the flux. The current will therefore be as shown.

When the coil is as in Fig. 49·14 (*c*), the flux through it will be increasing; the induced current will tend to diminish this flux and therefore be as shown. [These directions may also be determined by Fleming's right-hand rule.]

Measurement of the Earth's Magnetic Field.—The earth inductor provides us with a ready means of measuring the dip at a point on the earth's surface. The coil is connected to a ballistic galvanometer [and a series resistance if necessary to limit the throw] and placed with its plane at right angles to the earth's horizontal field. The coil is rapidly rotated through half a complete turn and the throw σ_1 observed. The coil is then placed so that it is horizontal and the throw σ_2 noted. If ϕ is the angle of dip, we have

$$\tan \phi = \frac{H_v}{H_0} = \frac{\kappa\sigma_2 R}{2A} \div \frac{\kappa\sigma_1 R}{2A} = \frac{\sigma_2}{\sigma_1}.$$

The determination of the actual values of H_0 and H_v is a little more difficult since another equation containing κ and R must be obtained. A long solenoid, P, Fig. 49·15, is connected through a reversing key, K, to a battery, C, an ammeter, A, and an adjustable resistance. Over the centre of P there is wound a coil, S, of many turns of fine wire which is connected to the earth inductor EI and ballistic galvanometer, G. The galvanometer kick when the inductor is rotated in the usual manner is first obtained, and we have, if H_0 is being measured,

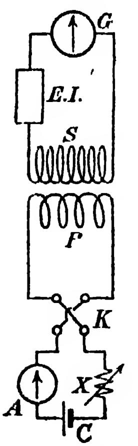


FIG. 49·15. — Measurement of H_0 and H_v with the aid of an Earth Inductor.

$$H_0 = \frac{\kappa\sigma_1 r}{2A}. \qquad (1)$$

The current in the primary is then established and adjusted until when it is reversed there is a galvanometer throw approximately equal to σ_1. Let it be σ_2. This throw is due to the fact that there is linked with each turn of the secondary $4\pi nia$ lines of magnetic induction, where a is the area of one turn—generally taken to be equal to that of the primary on which it is wound—and $4\pi ni$ is the field due to the current i in the primary which has n turns per unit length. When the current is reversed $8\pi nia$ is the change in the magnetic linkage with each turn of the secondary. If there are N turns in the secondary, the quantity of electricity passing through the galvanometer is $\frac{8\pi nNia}{r} = \kappa\sigma_2$. . . . (2)

From equations 1 and 2 we have

$$H_0 = \frac{4\pi n N i a}{A} \cdot \frac{\sigma_1}{\sigma_2} = \frac{2\pi n N I a}{5A} \cdot \frac{\sigma_1}{\sigma_2}$$

where the current, I, is measured in amperes.

Similarly H_v may be found.

Previously we have measured magnetic fields with the aid of magnets oscillating in them. The method just described possesses several advantages :—(*a*) Fields of all magnitudes may be measured, for if they are weak the galvanometer throw may be increased by using coils with a large effective area, whilst if they are strong this area must be reduced ; (*b*) with oscillating or deflecting magnets only the horizontal component of a field may be measured but the electromagnetic method enables the field in any direction to be measured.

Alternating Currents of Sine Wave-Form.—It has been shown that the current in an earth inductor at any instant is given by $i = -\frac{H_0 A}{r}\frac{d}{dt}(\cos\theta)$, i.e. $\frac{H_0 A\omega}{r}\sin\omega t$, where ω is the angular velocity of the coil and t the time measured from that instant when the coil passes through a position at right angles to the field. The current will therefore be zero when $\omega t = 0$, and reach its maximum value $\frac{HA\omega}{r}$ when ωt is $\pi/2$, i.e. when the coil is parallel to the field. The frequency of the current is $f = \frac{\omega}{2\pi}$. The expression for the current, which is termed an ***alternating-current*** may therefore be written $i = \frac{HA\omega}{r}.\sin 2\pi f t$. The quantity $2\pi f$ is denoted by ω. [To measure the frequency of an a.c. supply, cf. p. 622.]

The above expression for the instantaneous value of the current may be written $i = \hat{I}\sin\omega t$, where $\hat{I}$ is the maximum value of the current, and $\omega = 2\pi f$. If the time is measured from the instant when the coil makes an angle ψ with the lines of induction, the 't' in the above formula must be increased by the time required for the coil to rotate through an angle ψ, viz. $\frac{\psi}{\omega}$. Then

$$i = \hat{I}\sin\omega\left(t + \frac{\psi}{\omega}\right) = \hat{I}\sin(\omega t + \psi).$$

The angle ψ is termed the ***initial phase*** of the current. Alternating currents may also be represented by $i = \hat{I}\cos(\omega t - \phi)$; $-\phi$ is then initial phase.

Measurement of Alternating Currents.—Moving-coil ammeters cannot be used to measure an alternating current since the

moment of inertia of the coil about its axis of suspension is too large for the coil to follow the variations in current. Use must therefore be made of some effect which is independent of the direction of the current. The heating effect, being proportional to the square of the current, fulfils this condition and may therefore be used to measure alternating currents. The instrument must be calibrated by using direct currents of known value.

Let $i = \hat{I}\cos(\omega t - \phi)$ be the instantaneous value of the current flowing through a resistance R. Then in time, δt, the heat, δH, produced is given by

$$J.\delta H = R.i^2.\delta t.$$

In a cycle of period T, where $\omega T = 2\pi$, the heat generated will be

$$J.H = R\int_0^T i^2.dt = R(\hat{I})^2\int_0^T \cos(\omega t - \phi).dt$$

$$= R(\hat{I})^2\int_0^T\left[\frac{1 + \cos 2(\omega t - \phi)}{2}\right]dt = \tfrac{1}{2}R(\hat{I})^2.$$

Thus the heating effect is the same as if a direct current of value $\frac{\hat{I}}{\sqrt{2}}$ passed through R. This is called the ***root-mean-square (r.m.s.) current*** equivalent to the alternating current of maximum amplitude $\hat{I}$. The r.m.s. current is denoted by I and is given by

$$I = \frac{\hat{I}}{\sqrt{2}}.$$

The principle of a hot-wire ammeter has already been described [cf. p. 918]. In general the whole of the current to be measured is not passed through the hot-wire: a shunt is used. By removing the shunt and placing a high resistance in series with the wire, the instrument becomes a voltmeter, suitable for use on a.c. circuits.

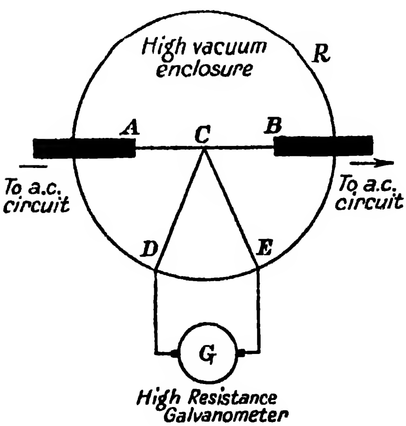


FIG. 49·16.

Very small alternating currents cannot be measured with an instrument of the type just indicated. A so-called vacuum-junction must be used—it is of special value when the frequency of the a.c. is high. It consists

of a thermocouple DCE, Fig. 49·16, with one junction C in close proximity to a fine wire AB carrying the current to be measured: it is insulated electrically from it. The junction and the hot wire are enclosed in a highly exhausted glass bulb, R, which prevents loss of heat by conduction and convection and so increases the sensitivity of the arrangement. The thermocouple is in series with a sensitive galvanometer, G, whose deflexion is observed. This deflexion is interpreted in terms of the initial value of the current by standardizing the instrument with small direct currents of known magnitude.

A Simple Test for Alternating Current.—If it is necessary to discover the nature of the current in the mains a very simple test is as follows:—a carbon filament lamp is connected to the mains and held between the poles of a horse-shoe magnet. If the current is alternating the filament oscillates very rapidly [if the magnet—i.e. field—is strong the filament may be fractured]. No such oscillatory movement is observed if direct current is used to light the lamp.

Faraday's Dynamo.—The simplest dynamo, a machine for the conversion of mechanical energy into electrical energy, for the production of direct current was first described by Faraday—it consisted of a copper wheel rotating in a uniform magnetic field normal to its plane and parallel to the axis of rotation. The apparatus and method have already been mentioned [cf. p. 946], but it suffers from the disadvantage that only small potential differences can be obtained at speeds which are practically possible.

The Principle of an Alternator.—A coil, PQRS, Fig. 49·17 (*a*), normal to a uniform magnetic field rotates about its horizontal axis with constant angular velocity. The ends of the coil are

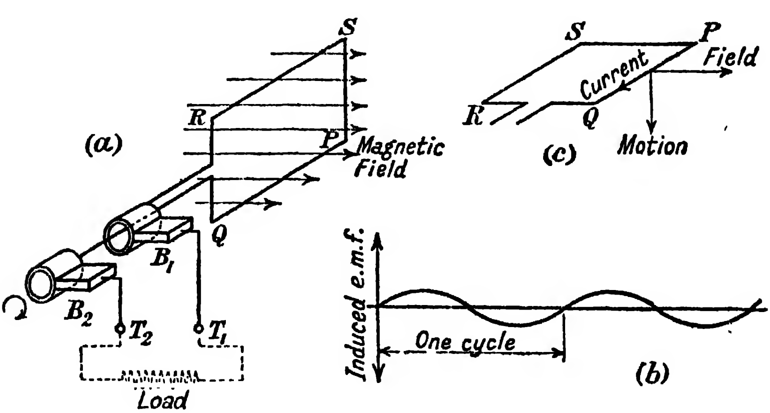


FIG. 49·17.—Principle of an Alternator.

joined to two metal rings each touching a carbon 'brush,' i.e. a carbon plate against which the ring slides. It has already been shown that the instantaneous value of the e.m.f. in this circuit is given by

$$-\mathrm{HA}\frac{d}{dt}\cos\theta = \mathrm{HA}\omega\sin\theta, \quad \text{since } \omega = \frac{d\theta}{dt}.$$

The e.m.f. therefore alternates between extreme values $+\mathrm{HA}\omega$ and $-\mathrm{HA}\omega$, and becomes zero twice in each complete revolution of the coil. Such an e.m.f. is termed an alternating one—it is represented by the sine curve shown in Fig. 49·17 (*b*). The direction of the current at any instant is determined by applying Lenz's law [cf. p. 941], or the right-hand rule. When PQRS is parallel to the field, the induced current flows from P to Q in PQ—cf. Fig. 49·14 (c); but when PQ occupies the position now occupied by SR, the current in it will be from Q to P, i.e. the current is an alternating one.

T_1 and T_2 are terminals by means of which the potentials at the brushes may be applied to an external circuit in which an alternating current then flows.

Rectification of an a.c.: The Generation of a d.c. from an a.c.—In order to produce a direct current from the alternating current obtained with the above apparatus a split commutator is

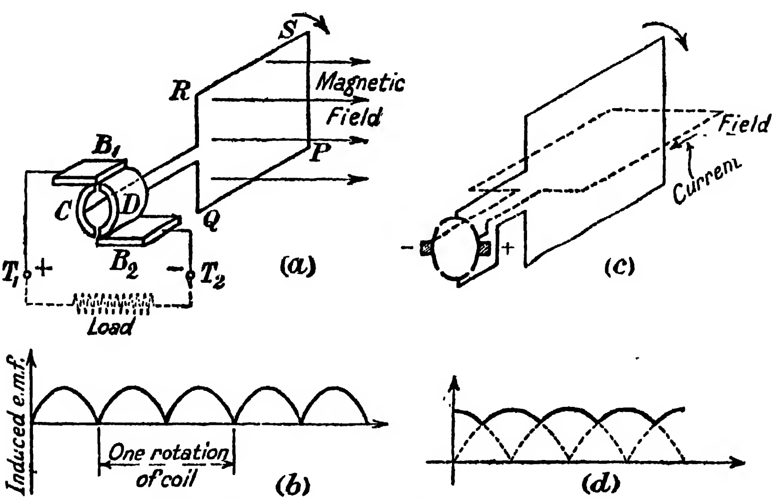


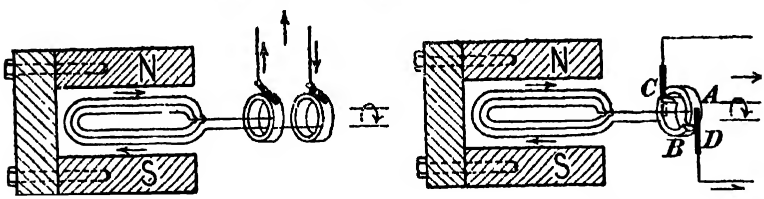
FIG. 49·18.—d.c. produced by Rectifying a.c.

used. This is shown in Fig.49·18 (*a*), and the apparatus is then known as a dynamo for the production of direct current. The carbon brushes B_1 and B_2 are alternately in contact with the

sections C and D of the split commutator as the latter revolves. The brushes are so placed that at the instant when the e.m.f. is changing its sign D leaves B_1 and makes contact with B_2. The current in the external circuit therefore flows in the same direction always, but it becomes zero twice in each revolution of the armature. Such a current is said to pulsate or to have a 'ripple'—it is represented graphically in Fig. 49·18 (*b*). These ripples may be smoothed by using a number of coils and a commutator with twice that number of segments. Fig. 49·18 (*c*) shows how two such coils may be arranged. It will be seen that the brushes are only in contact with two opposite segments in turn while the e.m.f. is in the neighbourhood of its maximum value, i.e. only the upper portions of the e.m.f. curves are effective. Fig. 49·18 (*d*) shows how the voltage across the terminals varies with the time.

Further Remarks about Dynamos.—Figs. 49·19 (*a*) and (*b*) show arrangements for the production of a.c. and d.c. respectively by rotating a coil in a magnetic field. The e.m.fs. will be greater than those obtainable with the simple dynamos hitherto considered, since the fields in which the coils rotate are increased by the use of permanent magnets as shown. But the field is no longer uniform, so that a steady rotation of the coil no longer produces an alternating current having a sine wave-form.

In actual machines the field in which the armature rotates is provided by a large electromagnet, excited by the current from



(*a*) (*b*)

FIG. 49·19.—Production of a.c. [non sine wave-form] and d.c.

the dynamo itself. When a dynamo is started up there is usually sufficient magnetism remaining in the magnet for a small current to be produced. If this current or a portion of it is allowed to flow through the coils of the electromagnet, the magnetic field increases. When the whole of the current generated flows through the coils of the electromagnet the machine is said to be ***series wound***. When only a portion of the current is used to excite the electromagnet the machine is termed a ***shunt-wound*** dynamo—cf. Fig. 49·20. The current from a series-wound dynamo varies considerably with the type of external circuit to which it is connected. If the resistance in this circuit is increased,

the magnetic field is reduced since the exciting current is reduced. With a shunt-wound dynamo, however, an increase in the resistance of the external circuit causes a larger fraction of the current to pass round the magnet's coils; this increases the field so that the current from such a machine tends to remain constant.

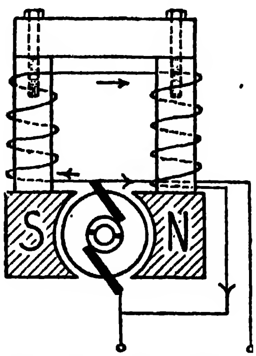


FIG. 49·20.—Shunt-Wound Dynamo.

The output from either type of machine is limited by the power available for driving it, for we must remember that as the current increases an augmented effort is necessary to turn the armature. This is due to the fact that the induced currents are in such a direction that they tend to stop the motion.

The potential difference between the brushes of a shunt-wound dynamo when the load is continually changing—as in a lighting circuit—changes slightly but nevertheless sufficiently to render this machine unsuitable for purpose of lighting. The defect is remedied by using a compound-wound dynamo, i.e. a shunt-wound machine in which a few series windings have been introduced.

The Gramme Armature.—The arrangements hitherto described for the conversion of mechanical energy into electrical energy have been chosen on account of their simplicity. The currents produced are very weak: they could be increased by augmenting the angular velocity of the armature but in practice this is impossible, partly on account of excessive wear in the running parts. In order to

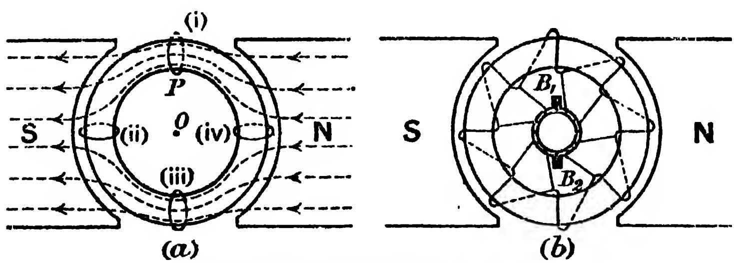


FIG. 49·21.—A Gramme Armature.

increase the output it is necessary to use iron to increase the variation of magnetic flux through the rotating coil. Only the essential features of one such arrangement, due to GRAMME, will be mentioned. A circular iron ring is placed between the semi-cylindrical pole pieces of an electromagnet N and S, Fig. 49·21 (a). The lines of magnetic induction tend to pass through the iron ring as shown.

Suppose that a closed coil, P, is fixed to the ring and that the ring rotates about an axis through O, the centre of the ring, perpendicular to the plane of the diagram. Then the magnetic flux through the coil is zero when it is in the positions (ii) or (iv), and a maximum when it is at (i) or (iii). It follows, therefore, that as the armature rotates an alternating e.m.f. will be produced in the coil. It will only be approximately simple harmonic in character.

If the coil is furnished with leads to a split commutator direct current may be obtained—it will, of course, pulsate.

To obtain a direct current suitable for ordinary use the armature is wound as in Fig. 49·21 (*b*). The ring is wound uniformly with copper wire and at regular intervals contacts are taken to separate segments of the commutator. The whole system comprising the ring, windings, tappings, and the commutator rotates with the axle to which the system is attached. The brushes are carbon plates fixed in space so that the different segments slide past each brush in turn once in each revolution. Leads attached to the brushes allow direct current to be taken from the generator to an external circuit. Let us see why the e.m.f. no longer pulsates.

If we consider any one of the segments of the commutator as it rotates, its potential will vary—in fact, twice in each revolution it will be zero and twice attain a maximum value (but with a different sign). The fixed brushes are arranged so that contact is made with each pair of opposite segments in turn when the potential difference between them is in the neighbourhood of its maximum value. If there is a large number of segments the potential difference between the brushes remains practically constant.

D.C. Motors.—It has already been shown [cf. p. 827] that a conductor carrying a current moves when placed in a magnetic field if the conductor is free. This is the basic principle of all electric motors. Any of the direct current dynamos just described will run as motors if connected to a suitable supply of direct current.

Back e.m.f. in Motors.—Suppose that R is the resistance of the armature windings and the coils of the field magnets. Let a battery of e.m.f. V be connected to the terminals of the machine. If the armature is at rest the current through it is given by $I = \frac{V}{R}$. Since R is small this current is large. But if the armature is allowed to rotate an induced e.m.f. will be set up in its windings. It will oppose the e.m.f. of the battery. This back e.m.f. (or counter e.m.f.) will increase as the armature rotates more quickly. Suppose that it is v at any instant. Then the current in the circuit at that instant is given by

$$I = \frac{V - v}{R}.$$

If the machine is frictionless and no external work is performed, the angular velocity of the armature will increase until $V = v$, i.e. no current is then being supplied by the battery. The armature would then revolve with constant velocity.

Suppose, however, the machine does work. If I is the current supplied by the battery, energy is then being dissipated at a rate VI watts, if the current and voltage are expressed in practical units. Now I is equal to $\frac{V - v}{R}$. The rate at which heat is developed in the armature windings is therefore

$$I^2R = \frac{(V - v)^2}{R} \text{ watt.}$$

Hence the rate at which external work is done is

$$\left[VI - \frac{(V - v)^2}{R}\right] \text{watt.} = \frac{V(V - v)}{R} - \frac{(V - v)^2}{R}$$

$$= \frac{V - v}{R} \,.\, v = W \text{ watt. (say).}$$

This expression shows that the external work done is zero if $V = v$, or if $v = 0$. The first condition has already been discussed: the second applies when the armature is at rest.

To find the condition that the power developed should be a maximum we obtain $\frac{dW}{dv}$ and equate it to zero, V and R being constant. We have

$$\frac{V - v}{R} - \frac{v}{R} = 0, \text{ or } v = \tfrac{1}{2}V,$$

i.e. the back e.m.f. is half the applied e.m.f.

It does not follow, however, that this is the most economical condition for running the motor, for

$$\frac{\text{Energy converted into useful work per second}}{\text{Energy supplied per second}} = \frac{\frac{V - v}{R} \,.\, v}{V\left(\frac{V - v}{R}\right)} = \frac{v}{V}.$$

If $v = \frac{1}{2}V$, the above ratio is 0·5. The speed at which a motor is run is generally such that $v = 0{\cdot}9V$, when the above ratio is 0·9.

Mutual Inductance.—It has been shown that whenever an electric current changes in a circuit (primary) there is established an induced electromotive force in any neighbouring closed circuit. The two circuits are said to possess ***mutual inductance.*** Let us examine this phenomenon more closely. It is well known that when a current flows in the primary circuit that there is present a magnetic field in the region round the circuit. Consequently there will be a definite number of lines of magnetic induction linked with any closed circuit situated in the magnetic field due to

the current in the primary. Now as long as the primary current remains constant and the positions of the circuits relative to one another do not change, the number of lines of magnetic induction linked with the secondary circuit is constant, and there is no induced electromotive force in that circuit. But if either the current in the primary, or the distance apart of the two circuits, varies, then there is at once produced an electromotive force in the secondary, and this continues to exist for as long as the number of lines of magnetic induction linked with the circuit is varying. Now the number of lines of induction linked with the secondary circuit depends on the strength of the current in the primary and on the relative positions of the two circuits, i.e. it depends on the current in the primary and on the geometry of the circuits.

If i is the current in the primary circuit, the number of linkages in the secondary is mi, where m depends only on the geometry of the system if no ferromagnetic material is present. When i varies, we have

$$\text{Electromotive force in the secondary} = -m\frac{di}{dt}.$$

The quantity m is termed the ***mutual inductance*** or ***the coefficient of mutual induction*** of the two circuits. It can be proved that for two given circuits m is constant; i.e. it does not matter which is the primary circuit and which is the secondary. The numerical value of the mutual inductance of a pair of coils is equal to the number of lines of magnetic induction linked with one of these coils when one e.m.u. of current flows in the other.

From the equation given above it is seen that the mutual inductance of a pair of circuits may be defined as

$$\frac{\text{the electromotive force in the secondary}}{\text{the rate of change of the current in the primary}}.$$

When the e.m.f. and the current are measured in absolute units of electromotive force and current respectively, the mutual inductance is measured in absolute units of inductance. The practical unit of inductance is the ***henry,*** and a pair of circuits has a mutual inductance of one henry when, if the current through one circuit is changing at the rate of one ampere per second, an electromotive force of one volt is induced in the other circuit. The value of a mutual inductance in practical units will be denoted by M.

Since 1 volt is 10^8 e.m.u. of potential difference, and 1 ampere is 10^{-1} e.m.u. of current,

$$\begin{aligned} 1 \text{ henry} &= \frac{1 \text{ volt}}{1 \text{ ampere . sec.}^{-1}} \\ &= \frac{10^8 \text{ e.m.u. of potential difference}}{10^{-1} \text{ e.m.u. of current . sec.}^{-1}} \\ &= 10^9 \text{ e.m.u. of inductance.} \end{aligned}$$

When Faraday constructed the apparatus shown in Fig. 49·1, he had really constructed the first mutual inductance. The mutual inductance of two coils is increased several hundred times when the primary is wound upon an iron core, but the mutual inductance is no longer a constant. Such an inductance is essentially a transformer if the primary is connected to an a.c. supply.

Practical Standard of Mutual Inductance.—A practical standard of mutual inductance—cf. Fig. 49·22—is constructed by winding a coil of many turns of wire (the secondary S_1S_2) on the middle portion of a long uniformly wound solenoid (the primary P_1P_2). Let this solenoid have n_1 turns per unit length, and suppose that the current through it is i. Then the magnetic field inside the solenoid is $4\pi n_1 i$ and is uniform over any cross-section near the middle of the coil. If the cross-sectional area of the coil is πr^2, then the flux through each turn in the primary is $4\pi n_1 i . \pi r^2$. Now outside the primary the magnetic field is very small so that if the secondary coil is very closely wound on the primary we are justified in stating that the flux through each

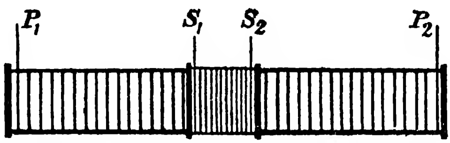


FIG. 49·22.—A Standard of Mutual Inductance.

turn of the secondary is also $4\pi^2 n_1 r^2 i$: if there are n_2 turns in the secondary, the total magnetic flux linked with that circuit is $4\pi^2 n_1 n_2 r^2 i$. If the current i is changing at a rate $\frac{di}{dt}$, the e.m.f. in the secondary circuit is equal to the rate at which the number of linkages through it diminishes, i.e.

$$-\frac{d}{dt}(4\pi^2 n_1 n_2 r^2 i) = -4\pi^2 n_1 n_2 r^2 . \frac{di}{dt}.$$

But this is equal to $-m . \frac{di}{dt}$, where m is the mutual inductance of the two circuits. Hence

$$m = 4\pi^2 n_1 n_2 r^2 \text{ absolute or e.m. units of inductance,}$$

or $$M = 4\pi^2 n_1 n_2 r^2 \times 10^{-9} \text{ henry.}$$

It cannot be emphasized too strongly that in an iron-free circuit the coefficient of mutual induction depends only on the geometry of the system, a fact which is at once apparent from the value of m deduced for the above pair of coils. Such an inductance is used to standardize the throws of a ballistic galvanometer [cf. Fig. 49·15, p. 952.]

Self Inductance.—When a current flows in a circuit there will be a definite number of linkages for that circuit due to the current in the circuit itself. This number depends only on the geometry of the circuit and the current, if no ferro-magnetic material is present. We may therefore write,

$$\text{No. of linkages} = \Psi = li,$$

where l is a constant for the circuit, and i the current in it. Now when the current changes there will be a variation in the number of linkages and consequently an induced electromotive force. In virtue of this there will be superimposed on the main current an induced current. By Lenz's law, the direction of this current is such that it tends to diminish an increasing current and to maintain one which is decaying, i.e. it tends to keep the magnetic flux linked with the circuit constant. It is owing to this fact that a current does not assume its final value at once; the interval of time may vary from a small fraction of a second to several minutes.

From Faraday's law of electromagnetic induction we may derive a value for the induced e.m.f. in a circuit due to changes in the current through it. We have

$$\text{Induced e.m.f.} = -\frac{d\Psi}{dt} = -\frac{d}{dt}(li) = -l\frac{di}{dt}$$

if l is a constant, i.e. if ferromagnetic materials are absent. The quantity l is termed the ***self-inductance*** or ***coefficient of self-induction*** of the circuit.

The units for self-inductance are the same as those of mutual inductance and a circuit has a self-inductance of one henry when, if the current is changing at the rate of one ampere per second, an opposing e.m.f. of one volt is set up in the circuit.

The coefficient of self-induction of a circuit is numerically the same as the number of lines of magnetic induction linked with it when the current through the circuit is 1 e.m.u. of current.

Effects of Inductance.—The presence of induction in a circuit or between a pair of circuits effects the current in several ways; it must be remembered, however, that the inductance has no effect as long as the strength of the current remains constant, but when it is growing or decaying the inductance plays an important part in determining the magnitude of the current at any instant. In all instances, the effects of induction are to oppose any variation in the current—when the current is increasing the inductance tends to make it increase more slowly; when the current is diminishing in intensity the inductance of the circuit tends to maintain it.

The presence of self-inductance in a circuit manifests itself by the so-called ***extra current*** appearing as a ***spark*** when an inductive circuit is broken. If the inductance is very big the spark is very bright and a person holding the ends of the wire which is fractured may receive a severe shock due to the high e.m.f. induced in the circuit. The formation of a spark is accounted for as follows. Initially the air in the gap is a non-conductor, but as the potential across it rises a stage is reached at which ionization sets in. The potential continues to rise and the ions become ionizing agents

themselves. The current increases and a spark passes. To avoid the effects of self-inductance in measuring resistances with a P.O. box, the coils are wound non-inductively, i.e. for each coil the free ends of the wire are soldered to the terminals, the wire made into a loop, and then wound on the bobbin.

Self-Inductance of a Solenoid.—Let us calculate the self-inductance of a long uniformly wound solenoid in air. The solenoid must be narrow compared with its length λ, so that we may neglect any want of uniformity in the magnetic field inside the solenoid. The magnetic intensity in such a coil is $4\pi ni$, where i is the current in e.m.u. and n the number of turns per cm. If r is the radius of the coil, the magnetic flux through each turn is $4\pi ni . \pi r^2$. The number of linkages associated with the coil is $4\pi^2 nr^2 i . n\lambda = 4\pi^2 n^2 r^2 \lambda i$. When the current varies the e.m.f. induced in the coil is

$$-l\frac{di}{dt} = -4\pi^2 n^2 r^2 \lambda . \frac{di}{dt}.$$

Hence $\quad l = 4\pi^2 n^2 r^2 \lambda$ e.m.u. of inductance

$\therefore \quad L = 4\pi^2 n^2 r^2 \lambda \times 10^{-9}$ henry,

where L denotes the self-inductance in practical units.

Faraday's Ring Transformer.—A transformer is an instrument whereby an alternating current supplied at one voltage may be changed to one at another—there is no change in frequency, however. If the voltage is raised and the current diminished we have

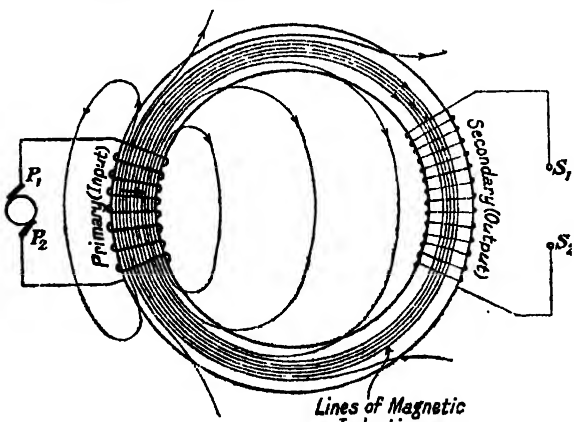


FIG. 49·23.—A Ring Transformer.

a step-up transformer: if the voltage is reduced, the current being increased, we have a step-down transformer. In each instance the principle is the same and we shall confine our attention to simple *transformers;* they were invented by FARADAY. In a step-up transformer the primary coil consists of a small number of turns of thick copper wire wound on an iron ring or core and insulated

from it and one another—cf. Fig. 49·23. The secondary then consists of a large number of turns of thin insulated wire. When an alternating current passes through the primary the iron is magnetized so that the lines of induction are first in one direction and then in the other; hence there is a continuous changing of the linkages with the secondary. The e.m.f. induced in the secondary is approximately n times that across the primary if n is the ratio of the number of turns in the secondary to that in the primary.

A very simple but inefficient transformer may be made as follows —cf. Fig. 49·24. A rod of soft iron, XY, is wound at one end with a coil, P_1P_2, consisting of several turns of thick wire. This is connected through a lamp, L (this acts as a resistance), to an a.c. supply. The other end of the rod is wound with a coil, S_1S_2, consisting of many turns of thin wire; the ends of this coil are connected to a telephone T. A note of the same frequency as that of the a.c. supply is heard in the phones.

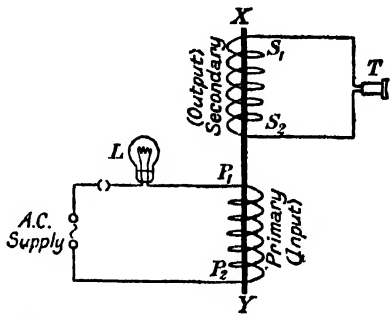


FIG. 49·24.—A Simple Transformer.

If no a.c. supply is available a battery may be used—a distinct click will be heard in the phones each time the battery circuit is made or broken, but when the current is established the phones are silent.

Further Notes on Transformers.—Transformers supplying electrical energy at 100,000 volts are in frequent use. In them good insulation of the turns of wire from each other and from the iron core is very essential. To satisfy these requirements all such transformers are immersed in oil, the dielectric strength of which is much greater than that of air.

The cores of transformers are not solid, but are laminated, i.e. they consist of strips of soft iron—or iron with a certain amount of silicon on account of the high permeability of this alloy—insulated from one another. In this way loss of energy due to the formation of large eddy currents is avoided; moreover, the temperature of the core does not become excessive when these currents are reduced in magnitude.

It is also important to select a material for the core in which the hysteresis loop is narrow, i.e. the loss of energy due to hysteresis in the core will be small.

For use with high-frequency currents, e.g. currents at radio

frequency, transformers cannot be wound on an iron core since the molecular magnets in the latter are unable to respond to variations in the magnetizing field by altering their orientation sufficiently rapidly. Such transformers have an 'air core' and the support for the wires must not contain any metal on account of the eddy currents which would be produced in them.

The Transmission of Electrical Energy.—Let G, Fig. 49·25 (*a*), be the generating station for the supply of electrical energy to a town T. Let R be the resistance of the leads from G to T and back. Let V be the voltage developed at the generator, and I the current to be delivered. Then the generator is developing energy at a rate of VI watts. The rate of loss of energy due to the Joule effect in the wires is I^2R watts. Hence the power available at T is $VI - I^2R = I(V - IR) = W$ (say).

From the above equation it appears that W will be practically equal to VI if I^2R is made small. One method of doing this is to make R small, i.e. the cable must have a large diameter. The

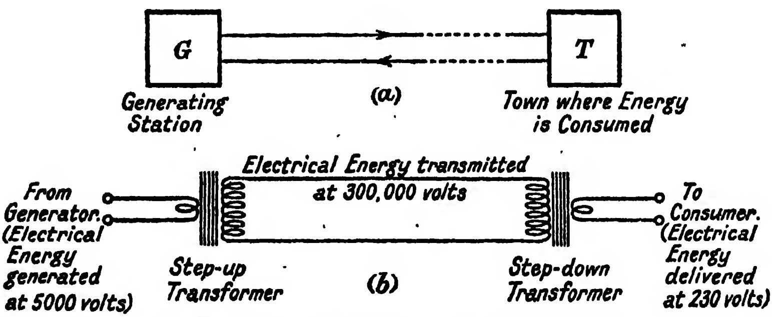


FIG. 49·25.—Transmission of Electrical Energy.[1]

initial outlay for a long cable of this type is prohibitive. Another method of reducing the losses due to heat developed in the wires is to make I small. But when this is done the voltage must be raised in order that the wattage supplied shall still be equal to W.

If V is to be large several considerations have to be noted. In the first place the insulation between the component parts of the generator would have to be almost perfect—practically this is impossible. Then again it is not desirable to supply current at a high voltage for domestic and factory use in general on account of the danger from shocks. Transformers supply the means whereby the current may be generated at a relatively low voltage, stepped up for the purposes of transmission, and then stepped down before being used by the consumer at a low voltage—240 volts is now the

[1] 240 V. since 1946.

standard voltage in Great Britain. Nowadays the energy is transmitted at 300,000 volts—cf. Fig. 49·25 (*b*).

Small Household Transformers.—For ringing electric bells current supplied from a 6-volt source is necessary. Instead of using cells a step-down transformer may be used if an a.c. supply at a

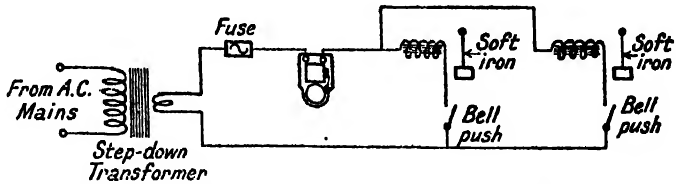


FIG. 49·26.—Transformer for Domestic Use—House Bell Circuit.

higher voltage is available. Fig. 49·26 shows the wiring of the necessary circuits when a bell may be operated from two different positions.

Example.—Give a diagram showing the essential features of a transformer suitable for supplying 4 V at 2 A and 1000 V at 100 mA when used on a 240-V a.c. main. Estimate a value for the current in the primary.

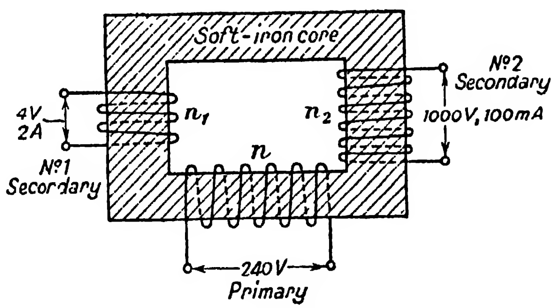


FIG. 49·27.

The scheme of wiring is shown in Fig. 49·27. If there is no loss of magnetic flux, i.e. all the lines of induction are confined to the iron core, then we may assume that

$$\frac{\text{E.m.f. in secondary}}{\text{No. of turns in secondary}} = \frac{\text{Applied p.d.}}{\text{No. of turns in primary}}.$$

Thus, with the notation indicated,

$$\frac{n_1}{4} = \frac{n}{240}, \text{ or } n_1 \simeq \frac{n}{60},$$

and

$$\frac{n_2}{1000} = \frac{n}{240}, \text{ or } n_2 \simeq 4n.$$

An estimate for the primary current I, can be made if it is assumed that

Rate of supply of energy in primary = sum of the rates of supply in both secondaries,

i.e. $$240I = (4 \times 2) + (1000 \times 0{\cdot}1) = 108.$$

$$\therefore\ I \simeq 0{\cdot}5 \text{ A}.$$

Energy Losses in Transformers.—The reason why the numerical answers in the above problem are given only approximately is that energy is lost when transformers are used to transfer energy from one circuit to another: this loss is accounted for as follows:—

(1) There is a leakage of magnetic flux, i.e. some of the lines of induction due to the current in the primary circuit do not pass entirely through the iron core but take shorter paths, partly in air, as indicated in Fig. 49·23.

(2) Loss of energy also occurs because eddy currents are produced in the iron core. These are diminished by using laminated cores, i.e. the core consists of a number of thin sheets of soft iron (or stalloy) laid together and insulated from one another by shellac varnish, tissue paper, or both. The effect of this arrangement is to cause the eddy currents to flow in high resistance circuits, so that the eddy current in each sheet is small.

(3) In addition, during each alternating current cycle the core is taken through a complete cycle of magnetization [cf. p. 984] and it is known that this entails a loss of energy proportional to the area of the cycle. This area is smaller for soft iron than for steel [cf. p. 984], and still smaller if an alloy of iron with 15 per cent. of silicon, known as stalloy, is used.

(4) Every transformer winding possesses resistance so that energy will be lost according to Joule's law [I^2R joule.sec.$^{-1}$]. The thicker the wire the smaller the loss of energy on this account.

The Induction Coil.—**The induction coil is a type of transformer whereby a comparatively strong current [direct] at a low voltage may be transformed into a weak current [intermittent] at a very high voltage. Essentially such a coil consists of a few turns of insulated thick copper wire, wound on a laminated core of soft-iron wire, surrounded by many turns of thin copper wire (double silk covered) on an ebonite or bakelite tube. This latter coil is termed the secondary. The core is not solid to reduce the magnitude of the eddy currents produced in it. When the current is switched on, and while it is reaching its final value, the iron becomes magnetized; during this process the effective magnetic flux across the secondary coil varies continuously so that a large p.d. is established between its ends, and if these are sufficiently close together an electric discharge takes place across the intervening space. If the primary**

current is then broken an e.m.f. (of opposite sign) is produced in the secondary circuit. The subsidiary adjuncts on a modern coil are merely there in order that the primary current may be repeatedly made and broken very rapidly. The manner whereby this is accomplished is indicated in Fig. 49·28 (*a*). The primary current is supplied from the battery B, and after traversing the upright and screw D it flows along the primary coil, P. The iron becomes excited and attracts the soft-iron hammer H, whereby the platinum contact A is withdrawn from D and the current is broken. The hammer, being supported at the extremity of a steel spring, and no longer attracted by the magnet, flies back, and the primary circuit is again closed. After each make and break the potential

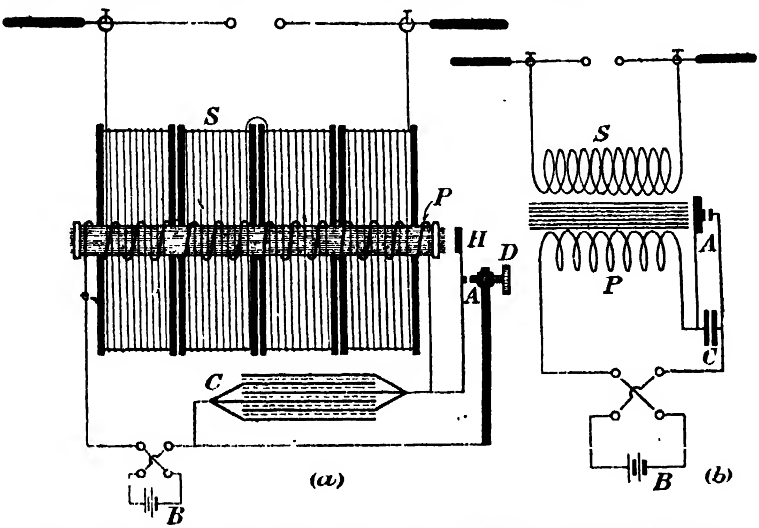


FIG. 49·28.—(*a*) An Induction Coil. (*b*) Diagrammatic Representation.

in the secondary assumes a very high value. It has been found that this arrangement works well if the sparking at A is reduced to a minimum. This is achieved by placing a condenser, C, in parallel with the spark gap. The condenser usually consists of sheets of tinfoil, paper being the dielectric, alternate sheets of the foil being connected together. The spark gap is furnished with platinum points, and since this is not easily vaporized the sparking is less persistent.

A diagrammatic representation of an induction coil is shown in Fig. 49·28 (*b*).

The relation between the primary current and the e.m.f. in the secondary of an induction coil is shown in Fig. 49·29. It will be noticed that the above e.m.f. is a maximum just after the primary

current has been broken. Where high voltages are essential this is the useful part of the e.m.f., and since it is large compared with the induced e.m.f. at other stages of the cycle of changes, it follows that the e.m.f. in the secondary is almost unidirectional but intermittent. A factor helping further to increase the secondary e.m.f.

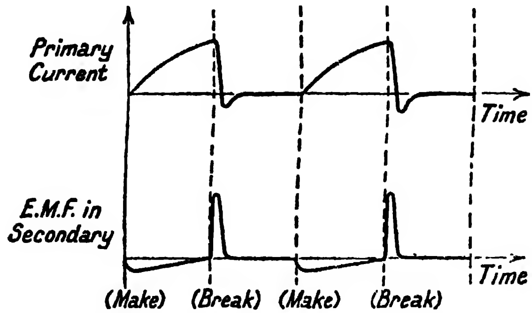


FIG. 49·29.—Relation between E.M.F. in Secondary and Current in Primary of an Induction Coil.

after the breaking of the primary current is the small reverse current in the primary due to the discharge of the condenser through it. Thus the condenser not only diminishes sparking at the 'make and break,' but also helps to increase the useful part of the secondary e.m.f.

Notes on the Construction of an Induction Coil.—In a modern large induction coil the main details of construction are as follows. The secondary coil is made up of a number of flat

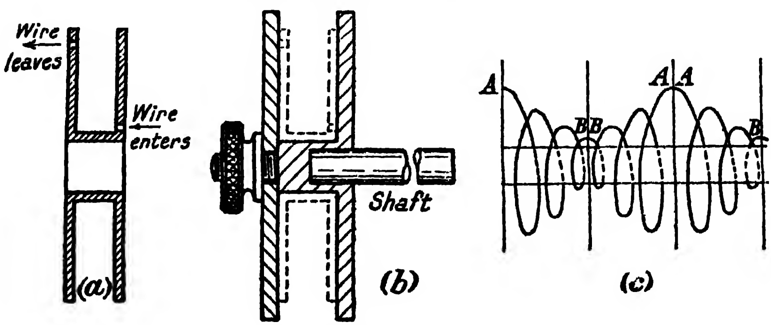


FIG. 49·30.—Construction of an Induction Coil.

sections 3 mm. to 5 mm. thick. Generally, these sections are wound separately and then assembled, the ends of the wire between adjacent sections being soldered together so that one continuous winding is formed. Each section is known as a 'pie.' To con-

struct one of these, a piece of ebonite is turned in a lathe until it has the shape shown in Fig. 49·30 (*a*). This is then supported on the 'former' shown in Fig. 49·30 (*b*), so that it may be wound with wire. While the winding is in process the wire is run through molten paraffin wax: this not only improves the insulation but helps to make each pie a solid entity.

The primary consists of a core of soft-iron wire mounted in a bakelite tube. The primary current is carried by a double layer of insulated thick copper wire wound on the above tube: the whole of this is mounted in another bakelite tube. The pies are then assembled on this as indicated in Fig. 49·30 (*c*). The soldering is done as each pie is added. When the requisite number of pies has been added, the whole is pressed firmly together, preferably while it is immersed in molten wax in a vacuum, so that all air bubbles shall be removed from between the wires.

The main advantage of constructing the secondary in this way is that high potential differences between neighbouring parts of the secondary are avoided—the risk of a breakdown of the insulation is thereby diminished.

The Coil-Ignition Set for a Motor-car.—B, Fig. 49·31, is the battery supplying current to the primary, P, of a small induction coil, when the spark gap is closed. A condenser is placed across this gap as in an ordinary induction coil. By means of a rotating shaft, a cam is driven so that the spark gap is opened and closed many times per second. In consequence of this a large p.d. is established between the ends of the secondary, S, of the coil, one of which

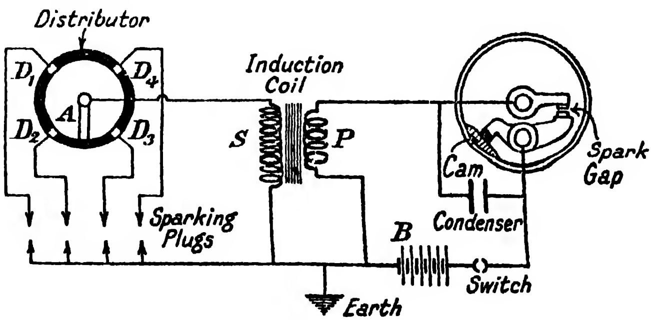


Fig. 49·31.—A Coil-Ignition Set for a Motor-car.

is earthed. The other end is connected to a rotating arm A, forming part of the so-called distributor. D_1, D_2, D_3, and D_4, are metal studs fixed in an insulating material. Each stud is connected to one end of a sparking plug, the other end being earthed. When, for example, A touches D_2, there is a large potential difference across the second plug and a spark passes igniting the petrol-air mixture in the second cylinder of the car. In turn the mixture in each cylinder is fired.

Wehnelt Electrolytic Interrupter.—In order that the circuit on an induction coil shall be broken rapidly so that the secondary e.m.f. may be as great as possible, the mechanical make and break described above is frequently replaced by a WEHNELT interrupter. A and B, Fig. 49·32, are two electrodes dipping into dilute sulphuric acid contained in a glass vessel. A is a large plate of lead connected to the negative pole of a battery, while B is a platinum wire, mounted in a porcelain sleeve, so that only its extreme point is in contact with the acid. The extent to which the platinum projects from its sheath may be controlled by a screw S. When the potential difference across this apparatus exceeds 30 volts the current density in the region of the electrolyte near the platinum point is very great and a considerable amount of heat is produced locally. Moreover, the amount of electrolyte decomposed here is large. A gas-layer covers the point of the platinum wire and the current is broken. A spark passes at this point —if necessary the energy of the spark may be increased by placing a coil possessing self-induction in the circuit—and this causes the bubble to detach itself from the wire. Contact is then re-established and the process repeated. With the aid of this interrupter, the circuit may be made and broken 1000 times per second. No condenser should be placed in parallel with the interrupter, for this diminishes the energy of the spark, thereby impairing the efficiency of the instrument.

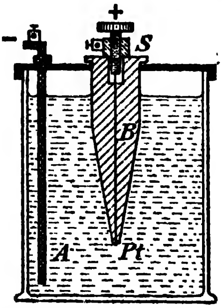


FIG. 49·32.—A Wehnelt Electrolytic Interrupter.

Mercury Interrupter.—An interrupter capable of making and breaking a circuit 1500 times per second but entirely different in principle from that designed by Wehnelt, is shown in Fig. 49·33 (*a*). A is a narrow metal tube whose upper end is made horizontal, while its lower end is connected to a metal sphere provided with several small holes, so that the mercury into which it is placed may pass freely towards the tube A. The whole is covered with an insulating oil and one which does not decompose when heated locally by a spark. B is a pulley by means of which the tube A may be rotated rapidly about a vertical axis. When this occurs a fine jet of mercury emerges from A and strikes C, a cylindrical 'mat' made of pieces of metal and ebonite arranged alternately—cf. Fig. 49·33 (*b*). These metal pieces are connected together and to one pole of a battery while B is joined to the other pole. Current will only pass from A to C when the mercury jet is in contact with one of the metal pieces. In this way a means of making and breaking the circuit rapidly is provided. A condenser placed across A and C diminishes the energy of the sparks occurring when the circuit breaks, i.e. the actual break takes place in a shorter interval of time.

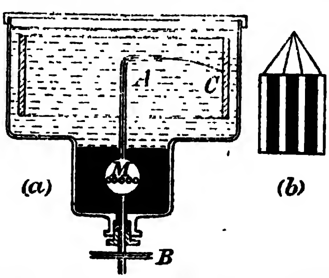


FIG. 49·33.—A Mercury Break (Mechanical Interrupter).

Telephony.—Graham Bell, about 1876, invented a telephone which first embodied the essential features of the modern instrument. It could be used both as a transmitter and as a receiver. The essential features of a modern form of telephone receiver are shown in Fig. 49·34 (*a*). A thin iron diaphragm D, varnished to prevent rusting, is situated behind the mouth-piece A, and immediately behind the diaphragm are soft iron pole pieces, P, fastened to the poles of a permanent horse-shoe magnet, M. This is held in position by the conical-shaped piece of brass C. Many turns of fine insulated

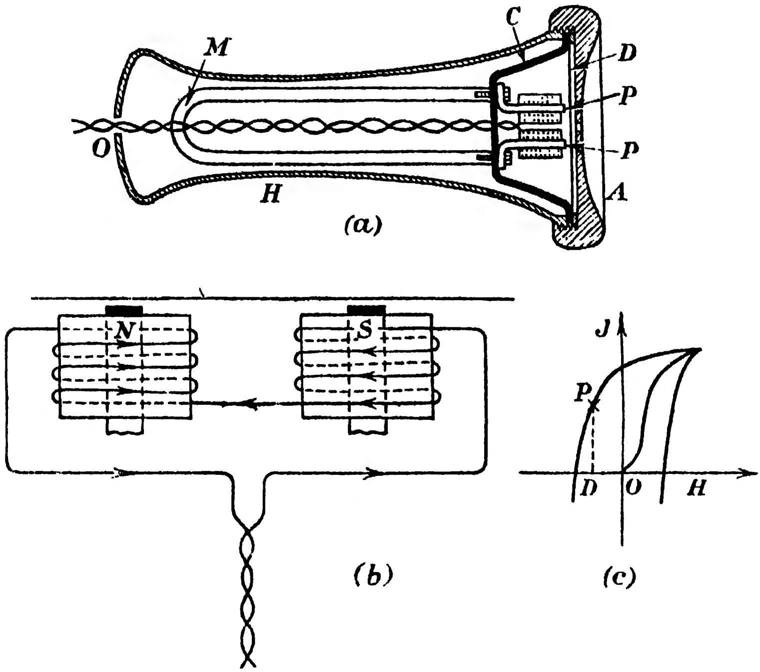


Fig. 49·34.—Telephone.

wire are wrapped round the pole pieces. These turns are arranged as in Fig. 49·34 (*b*) so that the pole strengths are each increased or decreased together, i.e. the pull on the diaphragm due to each pole is always increased or decreased. [The arrows in Fig. 49·34 (*b*) indicate the direction of the current when the numerical value of the strength of each pole is being increased.] H is an ebonite cover which serves as a holder. The leads from the magnetizing coils pass through a small aperture O in H. The permanent magnet M keeps the pole pieces magnetized in a condition corresponding to a point P on the steep portion of the hyteresis curve

shown in Fig. 49·34 (c). Hence any small change in the magnetic field causes a considerable variation in the pole strength of these pieces. If, therefore, rapid fluctuations of current, or intermittent unidirectional currents, are passed through the windings wound round the pole pieces, the diaphragm, whose natural frequency is somewhere in the range of audition is caused to vibrate, and if the frequency of the vibrations is within the limits of audibility, the sensation of sound will be produced. Moreover, it happens that the amplitude of the variations in current and of the displacement of the diaphragm are practically proportional to the amplitude of the soundwaves, i.e. there is little or no distortion.

To transmit a verbal message two such instruments were installed and connected by a wire and also to earth. No battery was used. When words are spoken into such an instrument the pressure fluctuations in the air cause the diaphragm to vibrate. There is then a piece of iron [the diaphragm] alternately approaching and receding from the poles of a magnet, so that the density of the number of lines in the magnetic field between the pole pieces undergoes corresponding fluctuations. Induced currents in the coils round P are thereby set up and these are transmitted to the second station. These cause identical variations in the field round the pole pieces of a similarly constructed instrument so that its diaphragm executes vibrations almost identical with those of the diaphragm at the transmitting station. This throws the air surrounding it into vibration and the sound is reproduced without distortion.

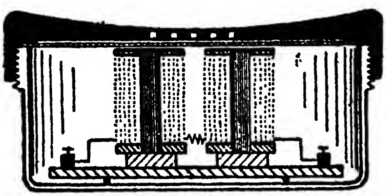


FIG. 49·35.—Earphone.

In Fig. 49·35 there is shown a section of an ear-phone. Its mode of action is similar to that of the telephone just described.

The Microphone.—Nowadays the above instruments are only used as receivers, for the currents transmitted when they are used for that purpose are too weak to enable the message to be conveyed across long distances. The microphone, invented by HUGHES, is an essential part of all modern transmitters. The following experiment enables us to understand the action of such transmitters :—

Experiment.—A pointed piece of carbon B, Fig. 49·36, rests in two notches cut in two carbon plates A, C. These are fixed in an insulating stand and connected to a telephone receiver R and a 4-volt battery. Any slight mechanical vibration causes the resistance at the carbon contacts to vary considerably and the current in the circuit will vary accordingly. If a watch is placed near to the instrument its ticks

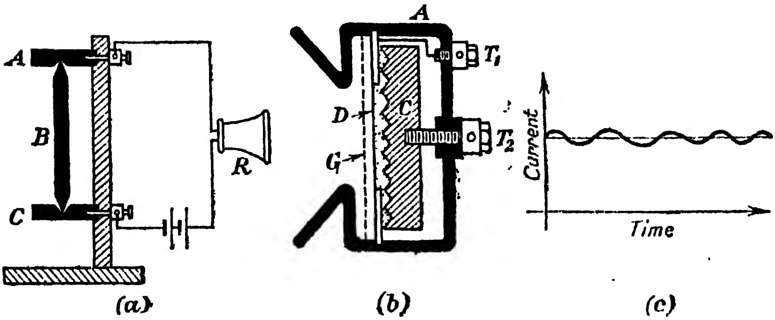


FIG. 49·36.—Microphone.

appear like the strokes of a blacksmith's hammer if the telephone included in the circuit is placed close to one's ear.

Modern Single-Button Carbon Microphones.—A simple carbon-granule microphone is shown in Fig. 49·36 (*b*). D is a thin carbon diaphragm, or rubber diaphragm coated on one side with powdered carbon. This is attached firmly at its edges to an outer ebonite case A to which a terminal T_1 is fixed: this terminal is connected to D. The case A also carries a second terminal T_2 which supports a carbon block C. The space between C and D is filled with loosely-packed carbon granules and under normal conditions these granules press lightly against one another. The terminals T_1 and T_2 are connected to an adjustable resistor, suitable battery, and ammeter, so that the steady current may be adjusted to the maker's stipulated value—this current is indicated in Fig. 49·36 (*c*) by the straight line parallel to the time axis. When sound waves are incident upon D it vibrates and the pressure between adjacent carbon particles alternately increases and decreases causing a corresponding increase and decrease of current through the circuit, i.e. the current is no longer steady but has superimposed upon it a fluctuating current whose frequency is identical with that of the incoming acoustic signal. If, for example, a tuning fork is held in front of a microphone, the total current varies as shown by the wavy curve in Fig. 49·36 (*c*), i.e. a sinusoidal current is superimposed on the normal steady current.

The Transmission of Speech.—In the preceding paragraph it has been shown how a microphone is used to produce electrical oscillations of the same frequency as the acoustical waves falling upon it. To transmit speech a microphone, M, Fig. 49·37, is arranged in series with a 4-volt battery, adjustable resistor and the primary P_1 of a step-up transformer, T_1. The secondary coil, S_1, of T_1 is connected [through the telephone exchange] to the secondary of the step-down transformer, T_2, the primary of which is connected

to a carbon-microphone M_2 and the necessary battery and resistor. Telephone receivers R_1 and R_2 are included at either end of the 'line' so that each station may transmit or receive signals. [The vibrations of the mocrophone's diaphragm at the transmitting end cause fluctuations in the value of the primary current so that a small current at a high voltage is produced in the secondary, and this is transmitted along the line and detected by the receiver at the other end of the system. It is necessary to use a return wire

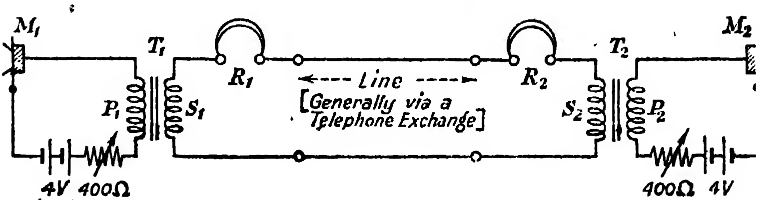


FIG. 49·37.

in this instance, for if it is dispensed with the inductive action of neighbouring circuits creates such a continual hum in the 'phones that the reception of speech becomes impossible.]

If transformers are not used the system will only be operative over short distances for, in other instances, the changes in resistance of the granules will be very small compared with the total resistance of the circuit which means that the fluctuations in current become too small to affect the diaphragm of the receiver.

EXAMPLES XLIX

1.—State and explain the meaning of Faraday's law of electromagnetic induction. Describe an induction coil and explain how it works.

2.—Describe and explain the action of (*a*) an induction coil or a transformer, (*b*) a microphone.

3.—Write a short essay on Faraday's discovery of electromagnetic induction.

4.—Describe the construction and explain the action of an alternating current transformer. Discuss its uses.

5.—A copper disc, 20 cm. in diameter, rotates on its axis normal to its plane 50 times a second in the earth's horizontal field. Assuming H_0 to be 0·18 oersted and the dip $\tan^{-1}$ 2·5, calculate the potential difference between the circumference of the disc and its centre.

6.—A circular copper disc of radius 20 cm. rotates on its axis 80 times a second. If the plane of the disc is normal to a magnetic field of strength 500 oersted, calculate the current when copper brushes touch the axis and periphery of the disc and the resistance of the external circuit is 10 ohms.

7.—Calculate the potential difference between the ends of the axle of a carriage wheel when a train is travelling at 60 kilometres per hour

across a horizontal plane where H_0 is 0·2 oersted and the dip 45°. The length of the axle is 1·2 metres.

8.—State Lenz's law of induction of currents. Calculate the maximum electromotive force induced in a circular coil of wire rotating 5 times per second about a diameter and at right angles to a magnetic field of strength 0·2 oersted, the effective area of the coils being 20,000 cm.^{2}.

9.—What is known concerning the electromotive force in a circuit placed in a magnetic field of varying intensity? A metal rod 1 metre long rotates about one end in a plane at right angles to a magnetic field of intensity 0·2 oersted. If the rod makes 2 revolutions per second, what is the difference in potential between its extremities?

10.—Under certain conditions a sphere of soft iron tends to move at right angles to the lines of force in a magnetic field. Describe the conditions and account for the phenomenon. (L. '23.)

11.—Describe and give the theory of the method by means of which the vertical component of the earth's magnetic field may be found.

12.—State Lenz's law of electromagnetic induction, and describe experiments to illustrate it.

Explain the essential features of a simple form of dynamo for producing 'direct' current.

13.—A coil of wire of 5 turns, each 1 cm.^{2} in effective area, is connected to a ballistic galvanometer of the moving magnet type, the total resistance in the circuit being 20 ohms. On introducing the coil into a strong magnetic field, the maximum deflexion recorded by the galvanometer is the same as that recorded when a condenser whose capacity is 5μF, and whose plates have a potential difference of 1 volt between them, is discharged through the galvanometer. What is the strength of the magnetic field? Would your result be correct if a moving coil galvanometer were employed?

14.—Define the absolute and the practical unit of inductance. What is the relation between them?

What is the self inductance of a solenoid of length λ cm. consisting of n turns of wire of cross-sectional area α cm.^{2}, the core on which the solenoid is wound having a permeability μ.

15.—Define mutual inductance.

Derive an expression in practical units for the mutual inductance of two coaxial coils of wire, the inner one of 250 turns of wire wound on a wooden core 50 cm. long and 2 cm. in diameter, while the outer one consists of 1,000 turns wound closely round the inner coil. Would you regard such an inductance as an absolute standard of mutual inductance?

16.—A circular disc 20 cm. in diameter rotates 5 times per second about an axis through its centre in a magnetic field normal to its plane. If the field has an intensity of 100 oersted, calculate the potential difference in volts between the axis and the periphery of the disc.

17.—A ballistic galvanometer, whose sensitivity is 200 divisions per micro-coulomb, is connected in series with a coil, A, and a resistance box which is adjusted to make the total resistance of the circuit 2,000 ohm. When a current is suddenly started in a neighbouring coil the galvanometer spot shows a throw of 95 divisions. Obtain a value for momentary change in effective flux (linkages) associated with the coil A.

Give a diagram to show how the quantity-sensitivity of the galvanometer can be determined if a standard cell and suitable condenser are available.

CHAPTER L

THE MAGNETIC PROPERTIES OF IRON AND STEEL

When a rod of soft iron is placed in a magnetic field the configuration of the field is changed for the lines of induction tend to concentrate themselves in the iron. In addition, north-seeking polarity is exhibited by the iron specimen at the end where the lines of induction leave it, and south-seeking polarity at the end where the lines of induction enter, i.e. the magnetic axis of the iron tends to point in the same direction as the magnetizing field. The iron has been ***magnetized by induction.***

Intensity of Magnetization.—It is assumed that a magnetic material consists of a large number of very small elementary magnets, all orientated at random when the substance is not magnetized. When placed in a magnetic field the orientation of the elementary magnets is disturbed, the tendency being for their axes to be parallel to that of the field. As this field is increased the number of magnets with their axes parallel to the field becomes greater, until finally all their axes are parallel and the substance is said to be magnetized to saturation.

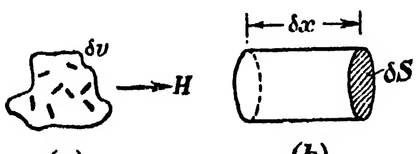


FIG. 50·1.

Let δv, Fig. 50·1 (*a*), be a small volume of a magnetic material, and let H be the magnetic field acting in the direction indicated. Now each elementary magnet has a definite magnetic moment: let each of these moments be resolved into components normal and parallel to H. If the small volume is large enough to contain a large number of elementary magnets the sum of the components normal to the field will be zero, since an initial random distribution was postulated. Let δM be the sum of the components parallel to the field. Then

$$\lim_{\delta v \to 0} \frac{\delta \mathrm{M}}{\delta v}$$

is called the ***intensity of magnetization***, J, for the material under the conditions considered.[1]

Now let Fig. 50·1 (*b*) represent a small cylindrical element of a magnetic material with its axis parallel to H, and its dimensions as indicated. If J is the intensity of magnetization, then

$$\delta M = J.\delta v = J.\delta S.\delta x.$$

Let $\pm\sigma$ be the surface density of the magnetism which must be placed on the ends of an equal small cylinder so that its magnetic moment shall be δM. Then its moment is $(\sigma.\delta S).\delta x$. Comparing these equations, we obtain $\sigma = J$.

Magnetic Intensity and Magnetic Induction.—The magnetic intensity at a point in air [strictly speaking, in a vacuum] has been defined as the force per unit positive pole on a small positive pole placed at the point. When it is desired to measure the force on such a pole inside a piece of iron, or other magnetizable substance, a cavity must first be made in the specimen so that the small pole may be introduced into it. Now the walls of the cavity will exhibit magnetic polarity which will, in general, contribute to the total force on the small pole in the cavity. This contribution will be determined, in part at least, by the shape of the cavity, which must therefore be carefully specified if the physical interpretation of this force is to have a definite meaning.

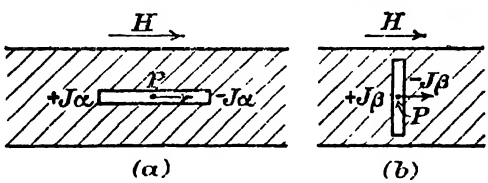


Fig. 50·2.

Let us consider the force per unit positive pole on a small positive pole, δm, at the point P, Fig. 50·2 (*a*), the centre of a cylindrical cavity, whose diameter is small compared with its length, and whose axis is in the direction of the magnetization at P. Then induced magnetism will appear on the ends of this cavity. If J is the intensity of magnetization, and α the cross-section of the cavity, the charges of magnetism at the ends of the cavity will be $J\alpha$ and $-J\alpha$, respectively. If $2l$ is the length of the cavity, the force on the small pole at P, due to the magnetism on the walls of the cavity, is $\left(\frac{J\alpha}{l^2} + \frac{J\alpha}{l^2}\right)\delta m$. This is zero, since the cavity is very long compared with its width. The force per unit

[1] Should it be necessary to distinguish this symbol from that for the mechanical equivalent of heat, we may use $J_0 = 4{\cdot}18$ joule.cal.$^{-1}$

positive pole at P is therefore due to the magnetizing field. Call it H.

Now consider the force on δm when this is at P the centre of a cavity whose length is small compared with its diameter—the cavity resembles a disc—Fig. 50·2 (b). Again let the axis of the cylinder be parallel to the field. Let β be the area of each plane face of the disc. It is only on these faces that induced magnetism will appear. Now the contribution to the force per unit positive pole on δm due to these induced charges of magnetism is 4πI, a result obtained from analogy with the corresponding problem in electrostatics [cf. p. 708].

The actual force per unit positive pole on the small pole in the cavity is obtained by adding together the two quantities H and 4πI. This total force per unit positive pole on the small pole is a measure of the ***magnetic induction***, B, of the material. Hence

$$B = H + 4\pi J.$$

The unit of magnetic induction is the ***gauss***.

Magnetic Susceptibility and Magnetic Permeability.—The quantity χ, defined by the equation $J = \chi H$, is termed the ***susceptibility*** of the material of the specimen.

The ***permeability***, μ, of the medium is defined by the equation $B = \mu H$. Since $B = H + 4\pi J$, it follows that

$$\mu = 1 + 4\pi\chi.$$

Strictly speaking, the equation $B = H + 4\pi J$, should be written $B = \mu_0 H + 4\pi J$, where $\mu_0 = 1$ is the permeability of a vacuum. The units for B and J are necessarily the same.

The Magnetic Permeability of Iron and Steel—Experimental Determination by a Magnetometer Method.—Let an iron rod of uniform cross-section be placed so that its axis is parallel to the lines of force of a magnetic field. The iron is magnetized by induction. Suppose that m is the strength of the induced poles, and $2l$ the distance between them. The magnetic moment of the rod is $2ml$. Hence, J, the intensity of magnetization in the rod is given by

$$J = \frac{2ml}{v} = \frac{m}{\alpha},$$

where v is the volume of the rod and α its cross-sectional area. Since the intensity of magnetization is not uniform throughout the rod, the above value for J must be regarded as that value which the intensity of magnetization would have if, while the product $2ml$ remains constant, the poles are considered to be at the ends of the rod. In other words this method measures m and we say that m is $J\alpha$.

In the relation $B = \mu H$, H is the intensity of the field causing the iron to become magnetized. Hitherto we have assumed that this has the same value as the field which would be present if the iron were removed. In general, this is not true, for the induced magnetic poles in the specimen react on the original field and produce what is termed a demagnetizing field. The effect of this field is negligible if the length of the specimen is 500 times greater than its diameter: hence very long thin wires must be used.

The experimental method then consists in subjecting such a wire to a uniform magnetizing field and determining the pole strength of the magnet produced. The uniform field is produced by passing a current through a long solenoid surrounding the wire, and in order that the wire may lie entirely within a uniform field, the length of the solenoid must be such that its ends project considerably

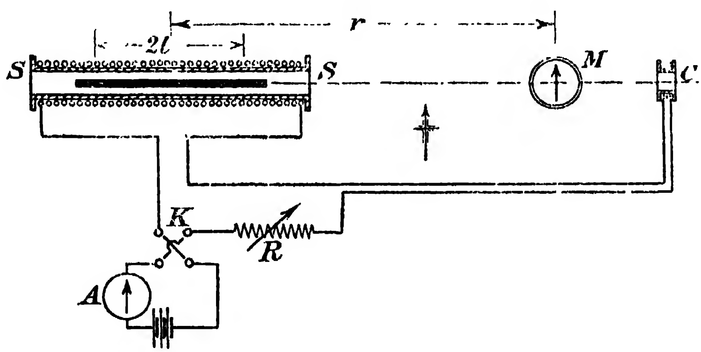


FIG. 50·3.—Measurement of Magnetic Permeability of a Ferromagnetic Substance.

beyond those of the specimen when the centres of this and of the solenoid coincide. The strength of an induced pole is measured by noting the deflexion of a magnetometer needle placed in a convenient position: a correction will be necessary for the effect of the more distant pole.

The specimen in the form of a rod about 40 cm. long and 1 mm. in diameter is placed centrally in a long solenoid SS, Fig. 50·3, so that when a current is passed through the coil the iron is in a uniform magnetic field. The axis of the solenoid must point east and west—then the horizontal component of the earth's magnetic field has no effect on the iron. The current through the solenoid is supplied from a battery of three or more cells, and its magnitude is controlled by means of a continuously adjustable carbon resistance R. An ammeter A measures the current and the direction of the current through SS may be changed with the aid of the reversing key K. The magnetometer M serves to

measure the pole strength of the iron when this is magnetized. The effect on the magnetometer of the field due to the solenoid itself may be compensated by means of a small coil C. This is placed in series with SS and its position adjusted so that the magnetometer deflexion is always zero when there is no iron in SS. This adjustment should be made with a large current in SS. To determine whether or not it is possible to obtain this adjustment when the apparatus has been set up, the coils S and C should be short-circuited in turn. Correct connections have been made if the deflexions given by M are opposite in the two instances. Initially the iron must be demagnetized by raising its temperature to that of a bright red heat or by subjecting it to an alternating magnetic field which is gradually reduced in strength to zero [cf. p. 984]. If the iron has been properly demagnetized the hysteresis curve finally obtained [cf. p. 984] will be symmetrical.

To commence the experiment a small current is passed through SS and corresponding readings of the current and the magnetometer deflexion observed. The current is increased step by step by adjusting R and the above observations made at each stage. This process is continued until the current reaches its maximum value. The permeability of the iron may be calculated as follows:—If J is the intensity of magnetization in the rod, α its cross-sectional area, and $2l$ its length, the magnetic moment of the rod is $2lJ\alpha$, since the intensity of magnetization is the magnetic moment per unit volume. If r is the distance from the centre of the rod to the magnetometer needle, H_1, the intensity of the magnetic field at M, is given by

$$H_1 = \frac{2(2lJ\alpha)r}{(r^2 - l^2)^2} = \frac{4J\alpha lr}{(r^2 - l^2)^2}$$

If θ is the angle of deflexion of the needle, and H_0 the horizontal component of the earth's magnetic field, $\tan\theta = \frac{H_1}{H_0}$. Hence

$$J = \frac{H_0 \tan\theta (r^2 - l^2)^2}{4\alpha lr}.$$

The strength of the magnetizing field due to a current I amperes in the solenoid is $H = 4\pi n\left(\frac{I}{10}\right)$, where n is the number of turns per unit length of the solenoid. Hence

$$B = 4\pi\left[n\left(\frac{I}{10}\right) + \frac{H_0 \tan\theta(r^2 - l^2)^2}{4\alpha lr}\right].$$

$$\therefore \mu = \frac{B}{H} = 1 + \frac{5}{2}\,\frac{H_0 \tan\theta\,.\,(r^2 - l^2)^2}{nI\alpha lr}.$$

The main advantage of this apparatus is that the specimen is not affected by the components H_0 and H_V of the earth's magnetic

field since its axis is normal to each of them. Unfortunately, however, the distance of either pole from the centre of the magnetometer cannot be determined easily for their exact location is unknown. This difficulty is avoided by placing the specimen in a vertical solenoid, but this necessitates that H_V should be compensated for otherwise the specimen is not completely demagnetized initially.

When corresponding values of B and H are plotted as in Fig. 50·4 the curve, OACD, obtained is termed ***a curve of magnetic induction.*** The dotted curve shows how the permeability varies with the field. We are not justified in drawing this curve in the neighbourhood of the origin since it is in this region that the experimental errors are large.

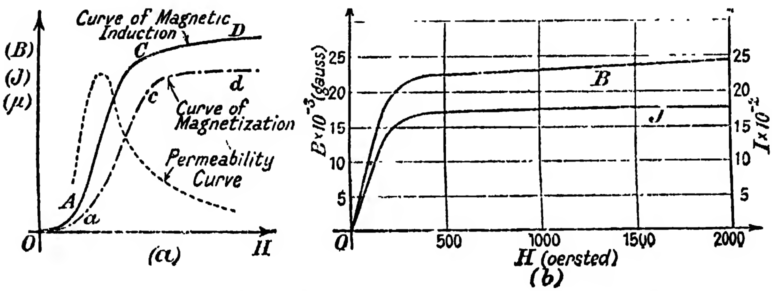


FIG. 50·4.—Curves of Magnetic Induction (B, H) and of Magnetization (J, H).

(a) At Relatively Low Values of H. (b) In a Strong Field.

A curve similar to the above is obtained by plotting J against H. This is known as the ***curve of magnetization*** [cf. *Oacd*, Fig. 50·4 (*a*)]. There is one very important difference between the two curves, however. When the magnetizing field H becomes large, the curve J–H tends to become parallel to the H-axis—cf. Fig. 50·4 (*b*), and when these conditions have been attained the iron is said to be ***saturated.*** The upper portion of the B–H curve never becomes parallel to the H-axis, for B is a function of H, viz., $B = H + 4\pi J$, and even if J were absolutely constant, B would still increase with H. It must be remembered, however, that B and H are plotted on very different scales, so that from the diagram it would appear as if B tended to reached a saturation value.

Hysteresis.—A more complete investigation of the behaviour of iron under the influence of a magnetic field may be obtained as follows:—When the current has reached its maximum value in the previous experiment the current is gradually reduced to zero

and the value of the current and the corresponding deflexion of the magnetometer are recorded at convenient intervals. It will be observed that even when the current is zero the magnetometer needle is still deflected, a fact showing that there is a certain amount of magnetism left in the rod. This is termed the ***remanent magnetism*** or ***remanence.***

If the current through the solenoid is then reversed, increased from zero to its maximum value and then reduced to zero, reversed again and increased to a maximum, the specimen will have been taken through a ***complete cycle of magnetic changes***. Fig. 50·5 shows how the intensity of magnetization varies with the magnetizing field for samples of iron and steel. Such curves are known as ***hysteresis curves,*** and the intercept OC on the J-axis measures

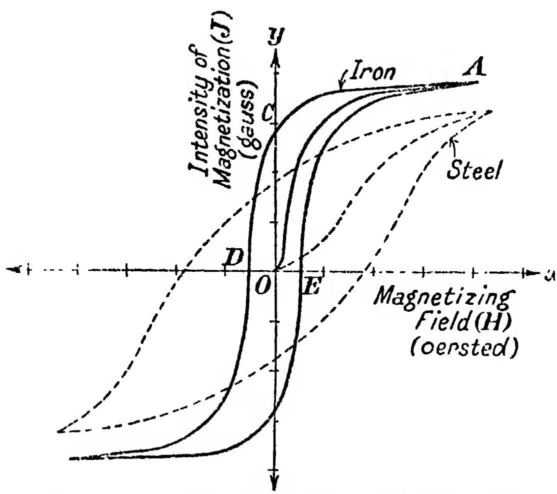


Fig. 50·5.—Hysteresis Curves for Iron and Steel.

the remanent magnetism in the specimen. This is called its ***retentivity***. The intercept OD on the negative H-axis measures that magnetic field which is necessary to reduce the remanent magnetism to zero. This field is termed the ***coercive force.***[1] The curves in Fig. 50·5 indicate that soft iron has greater retentivity than steel when both materials have been magnetized to saturation but that the coercive force for steel is greater than for iron. When a hysteresis curve is plotted to show how the magnetic induction varies during the cycle, the intercept on the y-axis is known as the ***remanent induction***. The negative intercept on the x-axis is not important for when $B = 0$, $J \neq 0$.

If a piece of iron is subjected to a series of hysteresis cycles, in which an initially large current is gradually reduced to zero, each

[1] or coercivity.

hysteresis curve gradually shrinks until finally the specimen is free from magnetism. This is most easily carried out with the aid of an alternating current—cf. p. 743.

Energy must be expended in order to magnetize a given specimen of iron or steel, but all the energy is not recoverable. When a magnetic material is taken through a complete cycle of magnetization, more energy is spent upon it than is returned to the source of energy (the battery). The difference appears as heat in the specimen. It may be shown that

$$\frac{1}{4\pi}\text{ (area of a hysteresis curve)}$$

represents the energy lost per cycle per unit volume of the material. The corresponding rise in temperature may only be 0·001° C., but if the specimen is subjected to an alternating field of, say, 50 cycle.sec.$^{-1}$, the rise will be 0·05° C. sec.$^{-1}$, or 3° C. min.$^{-1}$

Temperature and Magnetization.—When a steel magnet is warmed its pole strength and therefore its magnetic moment decreases. Its original strength is regained in part as the temperature resumes its initial value. Iron and steel are not attracted by another magnet when they are raised to a temperature above red heat. To show this an iron wire is made into a short coil and suspended near to one pole of a magnet. It is attracted by the magnet and finally rests in contact with it. A current is then passed through the iron wire so that it glows—it falls away from the magnet.

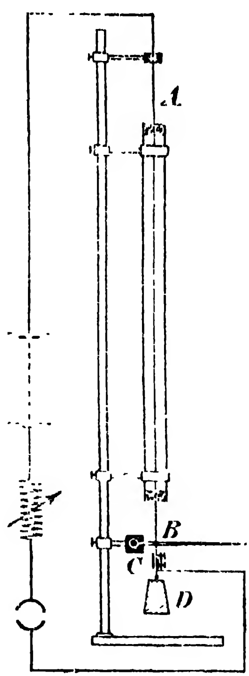


FIG. 50·6.—Experiment on Recalescence.

With pure iron it is found that at temperatures above 900° C. it ceases to be ferromagnetic but that it can be remagnetized when its temperature has fallen below this value. This temperature is termed the *recalescence point* for iron, for when a piece of iron is allowed to cool after being heated to a temperature greater than this then the wire glows suddenly as it passes through this temperature, i.e. heat is evolved in the specimen. Other abrupt changes occur in other properties of iron at this temperature, e.g. its resistivity changes suddenly. Hence an iron wire could

not be used as a resistance thermometer at high temperatures. Recent research has shown that at the recalescence temperature there is a sudden change in the arrangement of the atoms in the iron crystals. At temperatures below this point iron exists as α-iron, the atoms being arranged at the corners of cubes with other atoms at their centres, i.e. the atoms are arranged on a body-centred cubic lattice. At temperatures above the recalescence point iron exists as γ-iron and the atoms are arranged on a face-centred cubic lattice, i.e. atoms appear at the corners of the cubes and also at the centres of the faces of the cubes.

This rearrangement of the atoms is accompanied by a change in the length of the specimen which may be demonstrated as follows :—AB, Fig. 50·6, is an iron wire surrounded by a glass tube to protect it from air currents. It is supported from an insulated clamp and its lower end is attached to a pointer moving about a pivot C. A mass of lead, D, serves to keep the wire stretched. When a current is sent through the wire it is heated and its expansion is indicated by the downward motion of the pointer. When the recalescence temperature is reached there is a sudden contraction in length and the pointer moves upwards. As the temperature is increased beyond this point the length again increases. Opposite effects are noticed when the wire is allowed to cool. In addition to these opposite effects, when the recalescence point is attained—as revealed by the jerk in the motion of the pointer—the wire glows brightly for a few seconds. This is due to the large amount of heat liberated when the iron changes from γ-iron to α-iron.

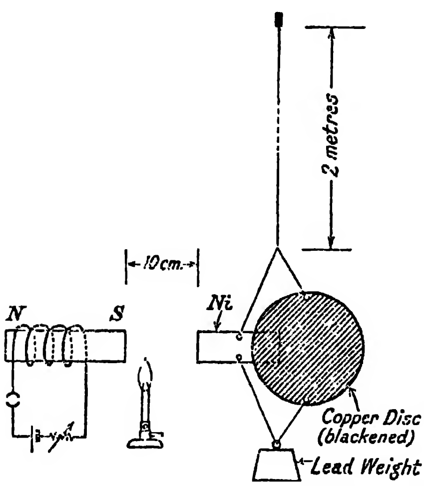


FIG. 50·7.—Shelford Bidwell's Pendulum.

Shelford Bidwell's Pendulum—' A Magnetic Engine.'—This is an interesting experiment, due to BIDWELL, based on the fact that nickel, a ferro-magnetic material, loses its magnetism when the temperature is above 340° C. —the critical temperature[1] for nickel. A nickel ' tongue,' 2 cm. × 1 cm., is soldered to a copper disc 3 cm. in diameter, the thickness of the materials being

[1] i.e. the Curie point, cf. p. 988.

about 2 mm. This forms the 'bob' of a simple pendulum of length 2 metres. If necessary, a lead weight may be attached to the system to keep the suspension stretched. NS is a straight electromagnet arranged in line with the nickel. The pendulum is set in motion, and the nickel would remain attached to the magnet were it not for the presence of the bunsen burner placed as indicated in Fig. 50·7. This heats the nickel, which loses its magnetism, and falls away from the magnet. The heat is conducted away from the nickel by the copper which is blackened so that this heat is quickly dissipated as radiant energy. By the time that the pendulum returns the nickel is below its Curie point temperature, it becomes magnetized as it approaches the magnet NS, and the process is repeated. [A long pendulum is selected so that the period shall be large.]

THEORIES OF MAGNETIZATION

Ferromagnetics, Paramagnetics and Diamagnetics.—Let us briefly recapitulate the main facts which are known concerning magnetism. Faraday discovered that many substances, and concluded that all, could be divided into two classes in so far as their behaviour in a magnetic field was concerned. He found that members of the first class, the so-called ***paramagnetics,*** when in the form of cylindrical specimens, arranged themselves with their axes parallel to the magnetic field in which they were freely suspended. On the other hand, members of the second class—the ***diamagnetics***—in similar form, set with their axes normal to the field. Amongst the paramagnetics, iron, cobalt, nickel, and certain of their alloys, were capable of becoming very strongly magnetized. Diamagnetism and paramagnetism are the general phenomena of magnetism, whereas ferromagnetism is a special case of paramagnetism. In passing, mention must be made of a series of alloys—the so-called ***Heusler*** alloys—which do not contain any iron at all, and yet exhibit marked ferromagnetic properties. One of these alloys contains 16 per cent. Al, 24 per cent. Mn, and 60 per cent. Cu.

Any satisfactory theory of magnetization must explain the following:—(i) The peculiar collection of properties known as ferromagnetism occurring in a few elements and their alloys.

(ii) The curve of magnetization, i.e. the curve exhibiting the relation between J, the intensity of magnetization, and H, the magnetic field. It must also account for the hysteresis loop. In the case of paramagnetics there is a linear relationship between H and J: for ferromagnetics, the relation is linear for low fields —this corresponds to the portion Oa of Fig. 50·4 (a); then there is a sharp rise, ac, and finally the approach to magnetic saturation —represented by the portion cd of the curve.

(iii) The 'residual magnetism' or permanent magnetism of ferromagnetics must also be explained. Paramagnetics do not show this phenomena. Modern work, at very low temperatures and with very strong magnetic fields, now indicates that under these conditions there is an approach by paramagnetics to saturation, i.e. the J–H curve is not linear. With ferromagnetics the condition of saturation is easily approached with relatively low magnetic fields, at ordinary temperatures. No paramagnetic substance showing residual magnetism or hysteresis is known.

(iv) The properties of a given specimen of iron depend on its past magnetic history.

(v) Magneto-striction, i.e. the change in the linear dimensions of a body occurring when it is magnetized.

(vi) In ferromagnetics, there is a temperature above which the residual magnetism disappears and the substance behaves like a paramagnetic body. This temperature is known as the ***Curie*** point.

(vii) The susceptibility, $\chi = J/H$, of a paramagnetic substance is inversely proportional to its absolute temperature, i.e.

$$\chi \propto \frac{1}{T}$$

For ferromagnetics above the Curie point, i.e. when they behave like paramagnetic bodies, the susceptibility is inversely proportional to $T - \theta$, where θ is the Curie point, i.e.

$$\chi \propto \frac{1}{T - \theta}.$$

The Molecular Theory of Ferromagnetization.—In this theory, due to WEBER, MAXWELL, and EWING, no postulate is made as to whether the elementary magnets considered are individual molecules or molecular aggregates. The demagnetization of a substance by heat or by rough mechanical treatment is explained by assuming that the elementary magnets are 'jostled' and rearrange themselves with their axes orientated at random in space. EWING has adapted his theory to make it fit experimental results quantitatively as well as qualitatively.

To account for the magnetic properties of iron and kindred materials it is assumed that the molecules of these substances are all small elementary magnets each having its own north-seeking and south-seeking poles. In an unmagnetized piece of iron, for example, these elementary magnets form closed chains for the elementary magnets that are near together must be arranged so that one pole of a molecular magnet is near to another of the opposite kind in a second molecular magnet. When all the molecular magnets form such closed chains the iron, as a whole, will exhibit no magnetic properties. When the iron is subjected to a magnetizing field each molecular magnet will experience a couple tending to rotate it so

that its axis is in the direction of the field. These couples will be balanced by the couples between the magnets which become operative when the molecular magnets are displaced from their zero positions.

To account for the main features of the magnetization curve *Oacd* shown in Fig. 50·4 (*a*) let us consider a group of four molecular magnets. Initially they will be arranged somewhat as in Fig. 50·8 (*a*). When a weak field is applied these magnets assume positions as in (*b*), and until the field is increased beyond the stage represented by Fig. 50·4 (*a*), the rotations of the magnets will be proportional to the field, i.e. the shape of the portion *Oa* of the curve is explained. As

(*a*) (*b*) (*c*) (*d*)

FIG. 50·8.—Molecular Magnets.

the field is still further increased a stage is soon reached in which the equilibrium of the magnets becomes unstable and they tend to arrange themselves with their axes in the direction of the field. This accounts for the sharp rise *ac* in the curve. On increasing the field beyond this only small changes are produced in the alignment of the molecules and the specimen reaches the stage of saturation—*cd*.

The above theory contains no suggestion as to the reason why the molecules are magnetic. Modern theory attributes the magnetic properties of ***all*** substances to the motions of the electrons in or among their constituent atoms. The magnetic effects caused by the motions of the positively charged nuclei are probably very small.

Theories of Paramagnetization and Diamagnetization.—The fact that the susceptibility of a paramagnetic substance is inversely proportional to its absolute temperature, whereas that of a diamagnetic substance is independent of the temperature, makes it clear that the phenomena must have different causes. Experimental results lead to the following conclusions.

Diamagnetism is a fundamental property of all substances, whereas paramagnetism is a possible feature, and when it occurs its effects are superposed on those attributable to the diamagnetism in the substance. Paramagnetic effects are very much larger than diamagnetic ones, so that the latter are masked considerably. Suppose that a body is strongly diamagnetic and that its paramagnetic properties are weak: the body will appear to be diamagnetic. When the temperature is raised, the paramagnetism disappears and only diamagnetism is left. Actually the amount of diamagnetism is unchanged but it will appear to have increased. This is the generally accepted explanation of the fact that the susceptibility of some diamagnetic bodies does appear to vary with temperature.

The variation of the susceptibility, χ, among the elements shows a periodicity agreeing with that of the Periodic Table. This suggests that magnetism is an atomic phenomenon.

Suggested Theory of Paramagnetization.—Suppose the electrons move round the nucleus of an atom in circular orbits. Then each electron is equivalent to a circular current which may be replaced by its equivalent magnetic shell of magnetic moment M. The direction of the axis of the equivalent shell is normal to the plane of the orbit. Now, for most atoms, the resultant magnetic moment of the equivalent magnetic shells is zero, or else very small. There are some atoms, however, such as sodium, in which the size of one orbit is very much larger than any of the others, i.e. there is an electron with a large orbit, and hence the magnetic moment of such an atom will not be zero. Such atoms will tend to align themselves with the axes of their resultant magnetic moments in the direction of the applied magnetic field; the material as a whole will become magnetized. Such bodies are the paramagnetics.

Now suppose that we have a large number of such atoms the direction of the magnetic moment of each being parallel to that of the applied magnetic field. As the temperature is raised, the atoms will acquire considerably more kinetic energy and the above alignment will tend to be destroyed. This accounts, in a qualitative way, for the decrease of the susceptibility of a paramagnetic substance with rise in temperature. It also shows that any agent tending to alter the orientation of the atoms will have a marked effect on the magnetic state of the substance, e.g. mechanical shocks.

Suggested Theory of Diamagnetization.—The electron theory, and also experiment, shows that a conductor carrying a current in a magnetic field is acted upon by a mechanical force—if the magnetic field is normal to the conductor, the direction of the mechanical force is normal to them both, its sense being given by the left-hand rule.

Now an orbital electron of an atom may be regarded as an electric current—in an atom with many extra-nuclear electrons there are many such orbits. It may be shown that the effect of a magnetic field on such a system is as if all the charges, as well as moving in their orbits, also rotated about the direction of the field. The electron would then describe a 'rosette.' Such a superimposed rotation of electric charges will give rise to a magnetic moment in a direction normal to the plane of the superposed rotation, i.e. normal to H. The sign of the effect is such that the field inducing it is opposed (general phenomenon in electromagnetism). The axis of the induced magnetic moment will therefore be opposed to the direction of the magnetic field, and the substance will therefore appear diamagnetic. Such phenomena will be common to all substances, if we assume that all atoms contain electrons. Paramagnetism is only shown by atoms having a particular type of electron orbit present in them.

The above is only a somewhat crude picture, but it appears to show that magnetism is a derived quantity due to electrons in motion, and the natural unit of magnetism would appear to be the magnetic moment of an orbital electron. Such a unit is termed the ***magneton***.

Ferromagnetism is of an entirely different order and its occurrence is rare. It must be connected with some peculiar atomic constitution—probably an inter-atomic effect.

EXAMPLE L

1.—Distinguish between the terms *intensity of a magnetic field* and *intensity of magnetization*. How would you investigate the relation between the magnetizing force and the intensity of magnetization for a soft iron wire? What difference would you expect to obtain if a steel wire were used?

2.—A cavity of width 0·1 cm.^{2} in the direction of the field used to magnetize a ferromagnetic material is made within that material but the cavity is very large in a direction normal to its width. If 120 erg.unit-pole.$^{-1}$ is the work done per unit pole in transferring a small magnetic pole across the cavity when the strength of the magnetizing field is 20 oersted., calculate values for (*a*) the permeability, (*b*) the susceptibility of the material.

3.—A rod of soft iron, initially unmagnetized and 0·20 cm. in diameter, lies in an E.-W. direction. The rod is 1 metre long and is situated inside a uniformly wound solenoid consisting of one layer of insulated wire of 1,920 turns and its length is 120 cm. If a steady current of 0·1 amp. is sent through the coil, calculate a value for the pole strength of the magnet using the following data:—

H [Oersted]	B [gauss]	H [Oersted]	B [gauss]
0·5	2,000	2·0	10,000
1·0	7,000	5·0	13,000

4.—Three metres above a laboratory bench is a horizontal steel girder, 4 metres long and of cross-sectional area 50 cm.^{2}, lying in the magnetic meridian. Calculate a value for the apparent value of H_0 at a point vertically below the centre of the girder, if the susceptibility of steel may be taken as 60 e.m.u. and the horizontal component of the earth's magnetic field as 0·200 oersted.

5.—a thin iron wire 60 cm. long and 0·02 cm.^{2} in cross-section is placed vertically so that its upper pole is level with and 10 cm. due W. of a magnetometer. The needle is deflected 15°. Given that H_0, the horizontal component of the earth's magnetic field is 0·200 oersted. and that the dip is $\tan^{-1}$ 2·40, calculate a value for the induction in the iron, assuming that its magnetism is due entirely to the earth's magnetic field and that the effect of the lower pole on the magnetometer may be neglected.

6.—A magnet 12 cm. long and 0·5 cm.^{2} in cross-section has a magnetic moment 900 erg.oersted.$^{-1}$ Assuming that the magnetic length of the magnet is five-sixths its geometrical length, calculate a value for the intensity of magnetization of the material of the magnet.

7.—A steel rod 36 cm. long (magnetic length = 30 cm.) and 0·05 cm.^{2} in cross-section lies inside a long solenoid having 9 turns per cm. length. The axis of the rod lies in an E.-W. direction and is initially unmagnetized. A magnetic needle, 30 cm. from the centre of the rod and due east of it, is deflected 20° when the steady current through the solenoid is 0·5 amp. If H_0, the strength of the earth's horizontal magnetic field, is 0·20 oersted., calculate values for (*a*) the pole strength of the magnet, (*b*) the intensity of magnetization, (*c*) the magnetic induction and (*d*) the permeability of the steel.

[Assume that the field due to the solenoid has been compensated in the usual manner.]

CHAPTER LI

THE DISCHARGE OF ELECTRICITY THROUGH GASES; X-RAYS; RADIO-ACTIVITY

The Spark Discharge.—When a potential difference of 20,000 volts is established between two terminals separated by about 1 cm. in air a spark passes. The actual potential difference necessary for the spark to pass depends upon the pressure and nature of the gas, the shape of the terminals, and their distance apart. The discharge is facilitated if a sharp metal point is attached to one of the terminals. While the spark lasts there is a considerable increase in the electrical conductivity of the gas, and a large current may pass.

The Discharge of Electricity in Gases at Low Pressures.—If a p.d. of a few thousand volts is placed across two electrodes sealed into a glass tube about 1 metre long and 5 cm. in diameter, no visible effect is seen and a sensitive galvanometer placed in the circuit fails to detect any current. A side tube leading from the above tube to a pump enables the pressure in the apparatus to be lowered and when this reaches a pressure equal to that of several cm. of mercury a long thin zigzag spark passes between the electrodes. When the pressure is reduced to about 1 cm. of mercury this spark widens and the tube is nearly filled with a bright column of light extending from the anode, the colour depending on the nature of the gas in the tube. It is known as the ***positive column.*** When the pressure is about one half-millimetre of mercury the appearance of the discharge undergoes an entire change. The cathode is covered with a soft pale light known as the ***cathode glow.*** Just beyond it is a region where colour is practically absent; it is termed the ***Crookes' dark space,*** C, Fig. 51·1. Beyond this appears a luminous column known as the ***negative glow.*** Its position is not influenced by that of the anode so that if this lies in a side tube the negative column is not directed towards it but goes straight on. A second dark space, F, generally appears beyond this column; it is called the ***Faraday dark space.*** It separates the negative column from the positive one which has now receded from the cathode and exhibits many striations. If the cathode

and anode have small holes at their centres luminous effects will now be seen in the regions beyond them. These will be discussed later.

When the conditions are such that the effects just described are seen, the p.d. necessary to send a current through the tube is a

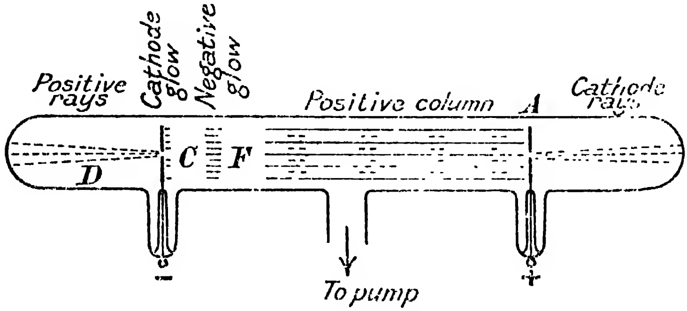


FIG. 51·1.—Discharge of Electricity through Rarefied Gases.

minimum. When the pressure is about 1×10^{-3} mm. of mercury the positive column disappears and a patch of green fluorescent light is observed on the glass wall opposite the cathode. It has been shown that this effect is due to something emitted from the cathode and travelling in straight lines along the tube. This latter fact may be demonstrated by placing a metal cross parallel to the cathode, when a sharply defined shadow is seen on the walls—

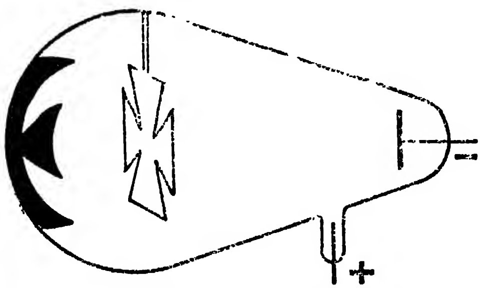


FIG. 51·2.—Mica-Cross.

cf. Fig. 51·2. If a magnet is placed near to the discharge tube the position of the shadow changes, showing that the path of this something in the tube has become curved. If an insulated metal plate is placed in the path of this something the plate acquires a negative charge. From these experiments one concludes that negatively charged particles travelling from the cathode must be responsible for these phenomena. They are electrons moving with a considerable velocity down the tube and are termed ***cathode rays***.

Cathode Rays.—The chief characteristics of these rays are as follows :—

(*a*) They travel in straight lines.

(*b*) When they strike soda glass they cause it to glow with a vivid green fluorescent light.

(*c*) When they are allowed to fall upon certain fluorescent materials many brilliant hues are produced.

(*d*) They can pass through thin sheets of aluminium foil without causing them to be punctured. LENARD first showed that this was possible and found that the rays were able to travel a few centimetres in air. The air glowed with a reddish violet light. The distance travelled by the rays in air could be determined with the aid of a fluorescent screen. When the screen was moved away from the aluminium window in the tube a point was reached when the screen failed to fluoresce. It was assumed that the cathode rays had then been brought to rest owing to frequent encounters with the air molecules in their path.

(*e*) The rays exert a mechanical force on the object with which they collide. The following experiment used to be cited in proof of this statement. If a small paddle with mica vanes is placed in the path of the rays and mounted on glass rails the paddle moves away from the cathode. This effect, however, is practically nothing more than a radiometer effect, for the actual force due to the electrons is less than one dyne.

(*f*) When the rays strike an object the temperature of the latter may rise considerably.

(*g*) They are deflected both by a magnetic and by an electrostatic field.

The Nature of Cathode Rays.—The fact that cathode rays were deflected by magnetic and electrostatic fields finally enabled the nature of these new rays to be established. For a long time one school of thought had maintained that they were a kind of invisible light travelling in waves through space, whereas another maintained that they were streams of negatively charged particles shot off from the cathode with considerable velocity. The pioneer work on this subject was carried out by SIR J. J. THOMSON who showed that they were negatively charged particles—or ***electrons.*** *Their mass was about $\frac{1}{1800}$th that of an atom of hydrogen and their velocity, depending, among other things, on the applied p.d. between the electrodes, ranged from one-thirtieth to one-third that of light, i.e. from 10^9 to 10^{10} cm. sec.$^{-1}$*

Fig. 51·3 is a diagram of an apparatus used by SIR J. J. THOMSON, to measure the velocity, v, and $\frac{e}{m}$, the ratio of the charge to the mass of an electron. This ratio is sometimes termed the ***specific***

charge of the electron. C is the cathode and A the anode, with a slit 1 mm. wide in it. Some of the cathode rays shot off from C when a suitable p.d. is applied to the tube which is filled with air at an appropriate pressure pass through A. The end S of the tube is covered with zinc sulphide which fluoresces at the point where the cathode rays strike it. Arranged on either side of the apparatus is a solenoid through which a current is passed. When these solenoids are excited the position of the bright spot on S changes. From the deflexion of this spot when the current through the solenoid is reversed, the ratio $\frac{mv}{e}$ is determinable when the strength of the magnetic field is known. H and K are two parallel plates in the tube and a known potential difference is applied to them. This electrostatic field also causes the spot of light to be displaced. It will be seen that the

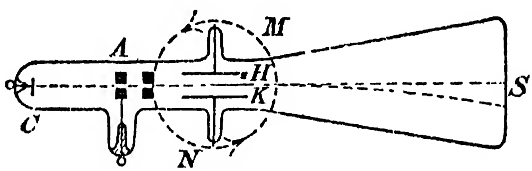


FIG. 51·3.—Apparatus for Determining e/m for Electrons.

electric and magnetic fields are at right angles to one another Whichever type of field is applied the deflexion is then in the same straight line. The directions and magnitudes of the fields are then selected so that the spot of light remains in its zero position when the fields are simultaneously applied. The velocity of the rays may then be determined. By combining these results, e/m may be calculated.

The dotted circle, M, with its anticlockwise current is merely a conventional way of showing that the north pole of the 'deflecting magnet' is above the plane of the paper.

Millikan's Method for the Determination of the Charge on an Electron.—The experimental arrangement used by MILLIKAN for the purpose of determining the charge on an electron is shown diagrammatically in Fig. 51·4. A cloud of very small oil droplets was produced in the chamber A containing very pure

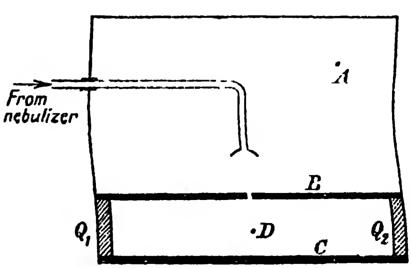


FIG. 51·4.—Millikan's Apparatus for Determining the Electronic Charge.

air. The drops were produced by means of a spray or nebulizer. These drops acquired a negative charge through friction and fell slowly downwards. A few of the drops found their way through a small opening in the base of the chamber. This base constituted the upper plate of a parallel plate condenser BC. These plates were horizontal, so that when they were charged the electric field at the centre was directed along the line of action of gravity. The above plates were insulated from one another by quartz rods Q_1 and Q_2 accurately ground so that the vertical distance between the plates was constant and the field at the centre of the condenser therefore uniform. A telescope enabled individual drops—such as D—to be observed when they were suitably illuminated by an arc lamp. This lamp was at a considerable distance from the apparatus so that the air between the plates of the condenser should not be heated, and any heating effect was further reduced by placing a water cell between the lamp and the experimental chamber. Moreover, to diminish the effect of convection currents in BC these were made negligibly small by placing the apparatus in a thermostat.

The telescope was provided with horizontal cross-wires and the time of fall of a droplet past these wires could be measured. The electric field was then suitably directed, and its magnitude chosen so that that the drop rose slowly: the time of transit across the wires was noted. The field could then be made zero (it is necessary to use a specially designed switch so that both plates of the condenser are automatically earthed when they are disconnected from the charging battery) and the experiment repeated. All this was done with one and the same drop and the charge on the drop could then be determined.

If in the course of the experiment the charge on the drop altered accidentally, or was altered by exposing the air between the plates to an ionizing agent—X-rays for example—the charge on the particular drop was always found to be an integral multiple of a certain elementary charge, e, which is the charge carried by one electron. The atomic nature of electricity was thereby established.

Millikan found that

$$e = -4{\cdot}774 \times 10^{-10} \text{ c.g.s. electrostatic unit of charge}$$
$$= -1{\cdot}591 \times 10^{-19} \text{ coulomb.}$$

[In passing we note that since the charge on a gram-ion is 96,450 coulombs, it follows that L, the number of atoms in a gram-atom, is given by

$$L = 6{\cdot}06 \times 10^{23}.$$

This is termed ***Loschmidt's*** or ***Avogadro's number.***]

Positive Rays.—If the cathode C, in Fig. 51·1, is drilled by a hole about 2 mm. in diameter rays will be seen in the region CD. These are termed ***canal*** or ***positive rays.*** They are deflected by electric and magnetic fields in directions contrary to those in which electrons are deflected. The magnitudes of the deflexions are much smaller for canal rays than for electrons. It has been shown that, in their simplest form, these rays consist of atoms which have lost an electron, i.e. atoms with a positive charge. By observing these deflexions Sir J. J. Thomson was able to calculate the masses of the atoms in the discharge tube. About 1919 ASTON improved this method of investigating the nature of a substance and showed that the values of the atomic weights of the elements as determined by chemical analysis were only average values. For example, chlorine with an atomic weight 35·46 as determined by chemical means, was found to be a mixture of two different chlorine atoms having atomic weights 35 and 37 respectively. Substances having different atomic weights but chemically indistinguishable from one another are termed ***isotopes.***

Röntgen Rays or X-rays.—In 1895 RÖNTGEN discovered that in addition to the green fluorescent light emitted from the point where glass was hit by cathode rays, this point was also the source of some invisible rays. These rays, unlike the cathode and positive rays, were not deflected by electric and magnetic fields. Moreover, he was unable to cause them to be diffracted or to produce interference effects. It was not until 1912 that SIR W. H. BRAGG and his son (now PROF. SIR W. L. BRAGG) showed that these rays did produce diffraction patterns when the structure of the diffraction grating was sufficiently fine. Crystals were the gratings they used. They proved that Röntgen rays were very short electromagnetic waves and therefore only differed from other light waves in the shortness of their wave-lengths.

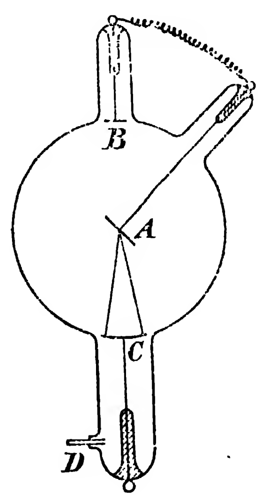


FIG. 51·5.—Gas-filled X-Ray Bulb.

The Gas-filled X-ray Tube.—An X-ray bulb of the type generally in use until about 1916 is shown in Fig. 51·5. The tube was exhausted until the pressure in it was about 3×10^{-4} mm. of mercury. C is the cathode, slightly convex, and A is the anode or ***anticathode*** as it is now generally termed. B is a second anode connected to the first. The exact part played by this electrode is not known and sometimes it is not fitted. When A and C are connected to

the terminals of a large induction coil so that C is at a negative potential with respect to A [the discharge from an induction coil is practically unidirectional] a beam of cathode rays converging upon the anticathode is produced. This becomes the seat of X-rays, which are produced whenever the motion of a swiftly moving electron is suddenly arrested.

When such a tube has been in use for some time it may fail to act. This is because the gas in the tube, which is essential for its operation, has become used up, i.e. the tube is 'hard.' It is 'softened' by heating with a bunsen flame the palladium tube, D, which, when hot, allows hydrogen to pass through into the bulb.

In 1913, COOLIDGE revolutionized X-ray technique by the introduction of a bulb fitted with a hot cathode. Fig. 51·6 shows a modern form of such a bulb furnished with two arms and exhausted as completely as possible. The one arm carries the wires conveying

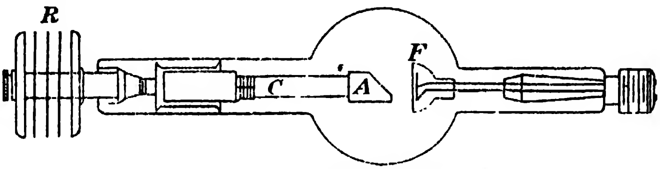


FIG. 51·6.—Coolidge X-Ray Tube.

the current necessary to heat a tungsten filament F. When this is heated to about 2,000° C. a copious supply of electrons is emitted. A few centimetres away from the filament is the anticathode A, inclined at 45° to the axis of the tube. The filament constitutes the cathode. It is surrounded by a hemispherical cap to focus the cathode rays on the target. When a large p.d. is applied to the tube the electrons are hurled with enormous velocity upon the anticathode where they are suddenly brought to rest and some of their energy emitted in the form of X-rays. A considerable amount of thermal energy is dissipated at the anticathode; this heat is conducted along the thick copper rod, C, supporting the anticathode, to the radiating fans R.

A transformer supplying a high alternating p.d. may be used with this latter tube which only allows the current to pass when A is positive with respect to F, i.e. the tube acts as its own rectifier.

Hard and Soft X-rays.—The penetrating power of the X-rays from a gas-filled X-ray bulb is greatest when the amount of gas in it is very small. The highly penetrating radiations from such a tube are often termed 'hard' X-rays; the rays from a tube in which the degree of vacuum is not so high, and which requires a lower potential difference across it to work, are termed 'soft' X-rays. In the tube with a hot cathode, the hardness of the rays

depends upon the potential difference across the tube; i.e. the higher the voltage the more penetrating is the radiation.

An X-ray Installation.—An installation for the working of an X-ray tube of the hot filament type is shown in Fig. 51·7. P is the primary of an induction coil, T_1 and T_2 being the terminals. These are connected to an a.c. source of supply, the current being controlled by a suitable resistance, or, more economically, by means of a choke. One lead from the secondary S is earthed and connected through a milliameter, MA, to the anticathode end of the tube. The other lead from the secondary is connected to the cathode. The cathode is a tungsten wire filament heated by the battery B. The filament current is controlled by a resistance R, and its value indicated by the ammeter A. Two pieces of metal, in the form of triangles, are connected to T_1 and T_2 and adjusted so that their points are about 1 mm. apart. If any high frequency e.m.fs. are induced in the primary a spark passes between these points owing to the high impedance of the

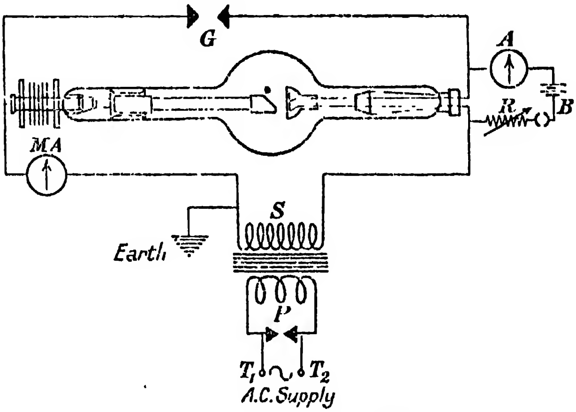


FIG. 51·7.—An X-ray Installation.

primary circuit for high-frequency currents. This prevents a possible breakdown in the primary circuit. With this arrangement no rectifying device is necessary since the tube acts as its own rectifier.

To protect the operator from X-rays the tube should be enclosed in a wooden box covered with lead at least 3 mm. thick. An aperture in the side of the box enables a portion of the rays to be used for experimental purposes.

Suppose that the filament ceases to emit electrons. The transformer is then supplying no current so that the potential difference across the tube rises. To prevent this from fracturing the tube a spark gap G is placed in parallel with the tube. The distance between its extremities is such that a discharge takes place across it when the potential difference applied to the tube tends to increase beyond its working limits.

Some Properties and Applications of Röntgen Rays.—X-rays render a photographic plate sensitive to a developer. If, therefore, an object, such as a human hand, is held between a

source of X-rays and a photographic plate, which is afterwards developed, a radiograph is obtained, i.e. there is produced an X-ray picture of the object. The X-rays are absorbed more by bone than by flesh, therefore the rays which traverse the bone are much reduced in intensity, so that their effects on the salts in the photographic plate are similarly diminished. If the subject which is under examination contains foreign metallic bodies, e.g. bullets, or coins and buttons which may have been swallowed, then the absorption of the X-rays is still more marked, so that the position of the metal can be located. In order to ascertain the depth at which such objects lie, two radiographs are taken at right angles to each other. If trouble is suspected along the alimentary canal, a large dose of bismuth is administered before a patient is examined. Bismuth compounds absorb Röntgen rays very easily, so that the alimentary canal, which contains the bismuth, stands in high relief against the rest of the picture. Nowadays, owing to the high cost of bismuth, barium sulphate is often used.

Some Further Uses of X-rays.—Röntgen rays find much application in detecting the presence of flaws, such as cracks and blow-holes, in metallic bodies. Where such faults occur the rays are transmitted more easily than elsewhere, so that a photograph reveals the defect easily. During the last few years, Röntgen radiation has been used to discover the arrangement of the atoms in crystals. Under certain conditions X-rays are reflected from the layers of atoms in the crystal so that, by measuring the angle of deflexion, the distance between the atoms can be calculated. In this way it has been shown that sodium chloride consists of small cubes at the corners of which the atoms are placed. Diamond and graphite are both allotropic modifications of the same element, but X-rays have shown that their structures are different. In diamond the carbon atoms are packed very closely together—hence its hardness; in graphite the atoms are so arranged that a certain plane of atoms is easily moved parallel to itself—hence the use of graphite as a lubricant.

The Ionization of air by X-Rays.—In general air does not conduct electricity. If this had not been so it would not have been possible to charge a body which was in contact with the air, nor for it to retain its charge for a considerable time. Towards the end of the last century it was noted that air, and other gases, could become conductors. In 1882 Giese first recognized that the conductivity was due to the fact that the gas atoms had become ionized, i.e. the creation of positive and negative ions in the air renders it a conductor. Gases, however, never ionize spontaneously—an external agent such as X-rays (or a radioactive material) is necessary.

Let us suppose that the air between two metal plates is ionized by X-rays and that the current between the plates is measured for different potential differences across them—cf. Fig. 38·6. A curve similar to that in Fig. 51·8 will be obtained. At small potential differences the ions move slowly towards the appropriate plate and if the X-rays continue to act there will be a current between the plates. But all the ions formed will not reach the plate for recombination occurs, i.e. neutral atoms are formed. When the potential is increased a stage is soon reached at which the ions move so quickly that only a negligible amount of recombination

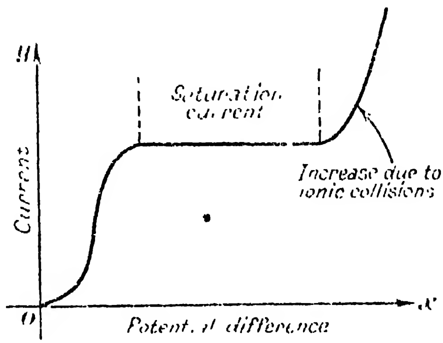


FIG. 51·8.

occurs. When this happens the current is a measure of the intensity of the ionizing agent for the number of ions reaching the electrodes in unit time is equal to the number of ions produced per unit time by the agent. Increasing the potential cannot increase the current, which has reached its saturation value. But if the increase in potential is considerable the ions may acquire sufficient velocity for them to create new ions by collision. When this occurs the current rises rapidly, the conductivity increases enormously and if the applied potential difference is sufficiently high a spark passes. The variation of current with potential when air is ionized by any ionizing agent does not obey Ohm's law.

Radium and Radio-active Substances.—The discovery of radioactive substances arose out of an attempt to find out whether naturally occurring substances emitted any penetrating radiations similar to that which had been discovered by Röntgen in 1895. Now the Röntgen or X-rays are characterized by the fact that they are able to penetrate considerable thicknesses of matter, ionize a gas, i.e. render it conducting, and cause luminescence on a fluorescent screen, both before and after passing through matter, although their efficiency in this respect is then considerably reduced. It is not surprising, therefore, to find that the substances which were

examined were those capable of glowing [i.e. phosphorescing] under the influence of light. Professor HENRI BECQUEREL placed a salt of uranium near to a photographic plate. After several hours a distinct mark was discerned on the developed plate—such a mark still persisted even when a thin silver screen was placed between the uranium and the plate, so that the darkening could not be due to any action between the silver salts in the film and any possible vapours emitted by the uranium. In addition, it was found that the radiation from uranium had properties similar to the above mentioned properties of X-rays. The CURIES suspected that the uranium they were using might contain a constituent far surpassing in activity that of the uranium itself. In 1898 MME CURIE isolated one of the salts of radium showing these remarkable properties to a very high degree, in fact its radiations were far more intense than those from the original mixture. It is now realized that radium is a by-product in the process of the emission of radiation from uranium. In 1903 RUTHERFORD and SODDY advanced the view that these phenomena must be attributed to the spontaneous disintegration of the atoms themselves; in this process new elements were formed which had properties different from the primary substances in which they had their origin. In the case of radium one of the products of the atomic explosion is the so-called ***alpha-particle***—this consists of a helium atom which has lost two electrons and which is hurled forth with a velocity approaching that of light. These particles possess an enormous amount of energy, so that when they impinge upon a screen of zinc sulphide the arrival of each alpha-particle manifests itself as a momentary flash upon the screen.

Soon after the discovery of these substances it was observed that the emitted radiations caused flesh to decay—it was hoped that such substances would be useful in checking and perhaps retarding some of the malignant growths to which mankind is subject. But whilst such experiments are still in progress and the results are promising, no definite conclusion has yet been reached.

Alpha-, Beta-, and Gamma-rays.—In 1899 RUTHERFORD showed that three distinct kinds of rays were emitted by radioactive substances. Fig. 51·9 shows schematically the action of a strong magnetic field normal to the plane of the paper on a narrow pencil of rays emitted from a small quantity of radium placed at the bottom of a narrow hole drilled in a block of lead. [The direction of the field is indicated in the conventional manner shown at the side of the diagram.] The α-rays are deviated slightly to the left while the β-rays suffer a much larger deviation to the right. The γ-rays are not influenced by the magnetic field. From the deviations thus produced we conclude that the α-rays carry

positive charges whilst the β-rays carry negative charges. From similar experiments made to discover the action of an electric field on these rays the ratio $\frac{e}{m}$ for the α and β rays has been determined. The α-rays are now known to be swiftly moving helium atoms which have lost two electrons, i.e. they are positively charged. These rays are able to pass through thin sheets of metal or glass, but they are completely absorbed by an aluminium plate 5 mm. thick. The β-rays are nothing else than electrons moving with enormous velocities, i.e. they are cathode rays moving with velocities much greater than those of the cathode rays we have previously studied. They are able to pass through the above aluminium plate. The γ-rays are now known to be very penetrating X-rays, i.e. their wave-lengths are about 100 times less than those from an X-ray bulb. These rays have remarkable penetrating powers and the intensity of the most penetrating γ-rays is only reduced to one-half by sheets of lead 1·4 cm. thick.

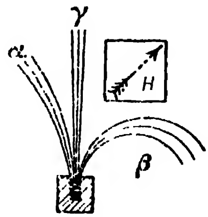


FIG. 51·9.—α-, β-, and γ-Rays in a Magnetic Field.

In the section of this book dealing with optics a short discussion of the Newtonian corpuscular theory and of Huyghens' wave theory of light was given. The former asserted that light consisted of a swarm of rapidly moving material particles, whereas Huyghens maintained that in optics we had to deal with the propagation of a state of motion. Now X-rays are a form of radiation having definite wave-lengths, whereas cathode rays are material particles moving with high velocities. In the instance of radioactive substances both types of radiation are found, i.e. some is corpuscular while the other is a form of light with a very short wave-length—far shorter than that of the rays emanating from an X-ray tube: even so, these radiations are grouped together under a common heading, for they transmit energy with great velocity through space.

A remarkable fact about radium is the enormous amount of energy which it emits as radiation. In the early days of radioactive discovery much speculation arose concerning radium as a possible perpetual source of energy, but the emission of radiation from a radioactive body is accompanied by a diminution in its mass, so that the ideal of the old alchemist remains as great an enigma as ever. A natural consequence of this large emission of energy from a radioactive material is that if a radium salt, contained in a tube, is placed in water, the temperature of the water rises. CALLENDAR succeeded in measuring the small heating effect due to 0·001 gm. of radium (in the form of radium chloride) by

means of an apparatus which he designed. The heating effect of the radioactive substance under examination is neutralized by the absorption of energy which occurs when an electric current is passed across the junction of two metals—the direction of the current must be such that energy is absorbed. The radioactive substance and the thermojunction were enclosed in a copper cylinder, and the electric current adjusted until there was no difference of temperature between the cylinder and an outer copper sphere. It was found that 1 gm. of radium emits energy at a rate of 130 cal. hr.$^{-1}$ This it continues to do for centuries.

The Nature of α-rays.—The direction of the deflexion of α-rays in a magnetic field shows that they are positively charged particles. The value of $\frac{e}{m}$ for these rays and their velocity may be obtained by compensating the deflexion of the rays in a magnetic field by an electric field arranged at right angles to the magnetic field and measuring the deflexion in either field alone as described on p. 995. For our present purpose it is convenient to take as our unit of electric charge that of a mono-valent positive ion, viz. $4{\cdot}77 \times 10^{-10}$ e.s.u., and as the unit mass that of a hydrogen atom, viz. $1{\cdot}66 \times 10^{-24}$ gm. In terms of these units $\frac{e}{m}$ for α-rays is $\frac{1}{2}$. Hence an α-ray is either a particle having a single elementary charge and an atomic weight 2, or else one with a charge 2 and a mass 4. Other possibilities naturally suggest themselves. Let us see how the problem was solved.

In an earlier section it has been mentioned that α-rays produce luminescence whenever they strike a fluorescent screen. If this screen is examined with the aid of a low-power microscope, it is found that the arrival of every α-particle is accompanied by a momentary flash of light, a fact indicating that these rays are discrete particles. These flashes of light are termed ***scintillations***, and by counting the number of these occurring in a measured time interval it is possible to determine the number of particles sent out from a speck of radium at the apex of a cone whose base is the screen on which the α-rays impinge. It is only the α-particles lying within this cone that are counted. The total number emitted per gramme of radium per second in all directions is easily deduced. If we are able to measure the total charge associated with the same number of particles, the charge on each particle follows at once. To measure the total charge on a known number of particles a minute source of the α-rays is placed in a vacuum and a charge collected in a given time by a metal plate, situated in the same vessel as the radioactive material, determined with the aid of an electrometer. Since an electrometer measures changes in potential

it is necessary to know the capacity of the instrument and its connections if the rate at which its charge changes is to be deduced from the rate at which the potential difference of the quadrants varies. The number of rays striking the plate is derived from the geometry of the system, etc. The charge on each particle was found to be $9{\cdot}548 \times 10^{-10}$ e.s.u. or $2e$, where e is the charge carried by a hydrogen ion in electrolysis, and $-e$ the charge on an electron. Since $\frac{e}{m}$ is equal to $\frac{1}{2}$, it follows that the mass of the α-particle must be 4, i.e. four times the mass of a hydrogen atom. It therefore seems that alpha particles might be ionized atoms of helium. To test this possibility, Rutherford allowed the alpha particles to penetrate through the walls of a very thin glass chamber into a discharge tube where the gas pressure was so low that no

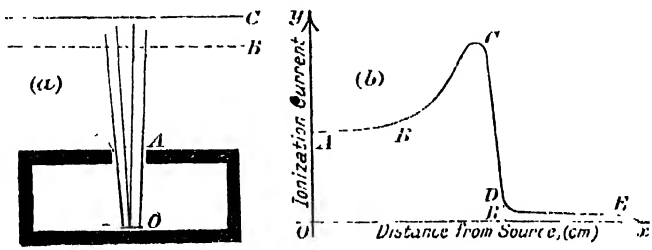


Fig. 51·10.—Range of α-Particles.

discharge could be passed through. At first no change occurred, i.e. no discharge of electricity took place when the electrodes were connected to the terminals of an induction coil. After a few hours sufficient α-particles had penetrated into the tube for a discharge to occur and the helium lines appeared when this was examined spectroscopically. The intensity of the lines in this spectrum increased as more α-particles penetrated into the discharge tube. The alpha particle is therefore an atom of helium which has lost two electrons, and therefore carries a charge $2e$. According to modern atomic theory an alpha particle is the nucleus of a helium atom, i.e. it is that part of a helium atom remaining when the latter has been deprived of its two outer electrons.

Alpha particles are characterized by the fact that after they have proceeded a certain distance in air they are no longer able to ionize the air. This phenomenon was discovered by Sir W. H. Bragg in 1904. His apparatus is shown diagrammatically in Fig. 51·10 (*a*). A layer of radioactive material is placed at O, a point on the bottom of a lead-lined box and some of the alpha particles emerge from a hole A in the lid of the box. B and C are the plates of a condenser, B consisting of a piece of wire gauze

so that the α-particles may pass into the condenser. By means of an electrometer the current between the condenser plates is measured. This is done for various distances of the condenser from the source. Since all the α-particles in the narrow pencil pass into the region between B and C, the variation of the current with distance is a direct measure of the ionization produced by the α-rays at different points of their path in air. The manner in which the current varied is indicated in Fig. 51·10 (*b*). At first the ionization current is almost independent of the distance. Then it increases to a maximum and finally falls very sharply almost to zero. If β and γ rays are present there is always a small residual ionization; this remains practically unchanged as the condenser is moved beyond the range of the α-particles under investigation. The distance OR is taken to be the range of the α-rays in air under the prevailing pressure conditions.

The α-particles emitted from a radioactive source have a very high energy content. When they collide with an atom they may knock out one of the extra-nuclear electrons of this atom. In doing this, the energy of the α-particle is decreased, but as it proceeds more and more collisions occur. A trail of positive and negative ions is therefore left behind, and the α-particle continues to operate in this manner until its velocity falls below a certain critical value when it is no longer able to eject an electron from an atom with which it may collide.

The Nature of β-rays.—From the deflexion of these rays in magnetic and electric fields it was soon apparent that β-rays were high-speed electrons. By means of methods already described $\frac{e}{m}$ and the velocity of these rays were determined. The values obtained for $\frac{e}{m}$, however, were not constant but the maximum value was the same as that for the electrons in a discharge tube across which the potential difference was not very great. The rays for which the specific charge was a maximum were the slowest. Now according to the theory of relativity the mass, m, of a body is related to its mass, m_0, when it is stationary—its so-called 'rest-mass'—by the equation

$$m = \frac{m_0}{\sqrt{1-\beta^2}},$$

where $\beta = \frac{v}{c}$ (c is the velocity of light in a vacuum). Hence

$$\frac{e}{m} = \frac{e}{m_0}\sqrt{1-\beta^2}.$$

Bucherer investigated the validity of this formula and his results were in agreement with it.

β-rays, unlike the α-rays, are not characterized by a definite range in air. This is because they are electrons and although their speed may be high their mass is small compared with that of an α-particle. When they ionize a gas atom by colliding with it they are deviated from their paths. The path of a β-ray in a gas at atmospheric pressure is therefore an irregular and devious one.

The Nature of γ-rays.—Since γ-rays are not deflected when subjected to the action of magnetic and electric fields they are regarded as being X-rays of very short wave-length. We have seen that X-rays are produced when swiftly moving electrons are stopped. It is believed that the γ-rays have their origin when a β-ray is stopped by the material responsible for its origin.

Helium from Radio-active Minerals.—The stop-cocks H and K of the apparatus shown in Fig. 51·11 are opened so that the apparatus is exhausted. The tube B containing a small quantity of freshly ignited charcoal is then cooled to liquid air temperature so that traces of gases and vapours still in the apparatus are absorbed. When this occurs an induction coil connected to the electrodes T_1 and T_2 fails to excite the tube C to luminescence. A small quantity of cleveite—a radio-active mineral - previously placed in A is then heated when gases are evolved. These pass over the charcoal which, at the temperature of liquid air, possesses the property of rapidly

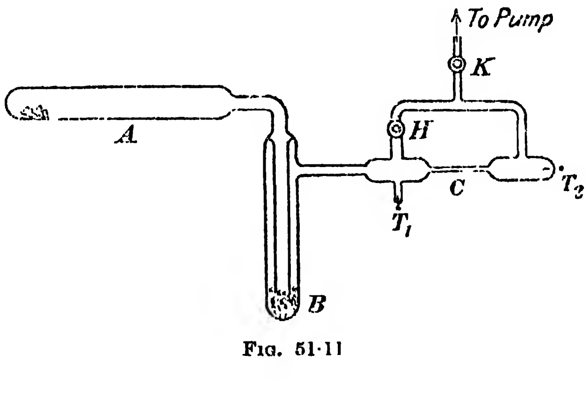


Fig. 51·11

absorbing such gases as air, carbon dioxide, etc. As the heating is continued and the induction coil kept in operation a faint yellow glow appears in the discharge tube. The intensity of this glow increases as the experiment continues. It is due to helium which, together with other gases, has been occluded in the mineral—the charcoal absorbs these other gases but not the helium. This helium is produced when the alpha particles ejected from the radio-active matter lose their positive charges and the normal helium atoms produced remain embedded in the cleveite. When the mineral is heated this helium, which is radio-active in origin, is expelled along with other gases which are always occluded in solids.

Cosmic Rays.—For a long time evidence has gradually been accumulating to show that at the earth's surface there is highly penetrating radiation very similar to γ-rays. For example, a gold-leaf electroscope slowly loses its charge even when the leak along the support of the leaves is prevented. The rate at which this charge is lost increases when the observations are made at high altitudes. This suggests that this radiation comes from space. MILLIKAN has recently carried out a number of experiments in this connection and he has found that the rate of leak of the electroscope from this cause decreases when it is sunk to different depths in lakes. The origin of these cosmic rays is uncertain.

Thermionics.—When certain inorganic salts are heated positive ions, i.e. atoms or molecules charged positively, are emitted. Thus impure aluminium phosphate emits sodium ions. Metal wires emit both positive and negative ions when heated, but at temperatures above 1,000° C. the emission consists almost entirely of electrons and are termed ***thermions.***

The Diode.—This consists of a highly exhausted glass vessel with a tungsten filament F and a collecting plate (or anode) P. A diode and its connections are shown in Fig. 51·12 (*a*). The filament is heated by current from a cell, C, controlled by an adjustable resistance, R. The negative end of the filament (by convention)

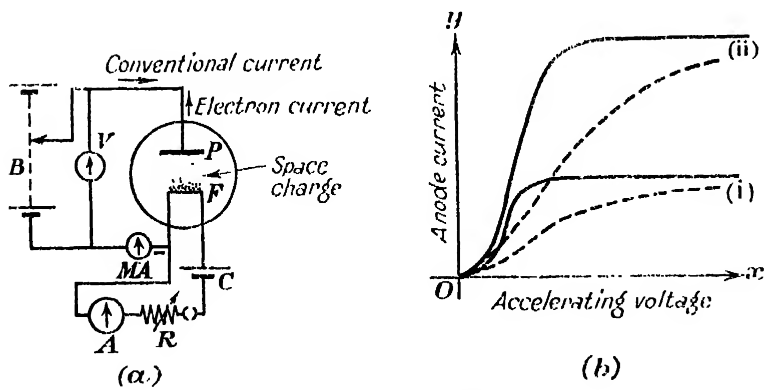


FIG. 51·12.—The Diode.

is connected through a milliammeter, MA, to the negative end of the battery B. The plate P is kept at a positive potential, measured by a very high resistance voltmeter V, with respect to the filament. When the temperature of the filament is sufficiently high electrons are emitted from it and most of these are collected at P. By varying the accelerating voltage, while the current through and

hence the temperature of the filament are maintained constant, and measuring the anode current, curves similar to (i), Fig. 51·12 (*b*), will be obtained. Curve (ii) is obtained when the filament temperature is increased. They constitute the ***characteristic curves*** of the diode. The first part of each curve obeys a $\frac{3}{2}$ power law, i.e. $I \propto V^{\frac{3}{2}}$, but eventually each becomes parallel to the x-axis showing that the current approaches a saturation value. This state of affairs is reached when all the electrons arrive at the collecting plate. It is most readily obtained when the plate P is cylindrical and entirely surrounds the filament. If, as frequently occurs, the filament is in the form of a loop extending beyond the ends of the cylindrical electrode, saturation is not so easily obtained. The dotted curves in the diagram are typical and it will be noted that the saturation current is approached more closely when the temperature of the filament is low. When the filament is hot, its tip, which is cooler than the rest since it is held in a support, is also emitting electrons and high voltages are required to bring all these to the plate.

It will be noted that the relation between voltage and current is not a linear one, so that Ohm's law is not valid in this instance.

If the plate P is connected to the negative end of the high-tension battery no plate current is obtained. It therefore follows that when such a tube is connected to a source of alternating current the current will only pass in one direction, i.e. the current has been rectified.

Example.—Electrons from a hot filament are shot across a vacuous space to a collecting plate maintained at a potential of + 100 volts relative to the filament. Assuming that the specific charge for electrons is $1{\cdot}77 \times 10^7$ e.m.u. gm.$^{-1}$, calculate the velocity of the electrons when they reach the plate.

$$\frac{e}{m} = 1{\cdot}77 \times 10^7 \text{ e.m.u. gm.}^{-1}$$

Now 1 e.m.u. of quantity $\equiv 3 \times 10^{10}$ e.s.u. of quantity.

$$\therefore \ \frac{e}{m} = 1{\cdot}77 \times 3 \times 10^{17} \text{ e.s.u. gm.}^{-1}$$

Also 100 volts $\equiv \frac{1}{3}$ e.s.u. of potential difference.

Let v be the velocity required. Then kinetic energy $= \frac{1}{2}mv^2$. But this is equal to the work done by the field on the electron, viz. eV ergs where e and V are in e.s.u. Hence

$$\tfrac{1}{2}mv^2 = eV.$$

$$\therefore \ v = \sqrt{\frac{2 \,.\, e \,.\, V}{m}} = \sqrt{3{\cdot}54 \times 10^{17}}$$

$$= 6 \times 10^8 \text{ cm. sec.}^{-1}.$$

[N.B.—$\beta = \dfrac{v}{c} = 0{\cdot}02$, i.e. the velocity of the above electrons is one-fiftieth that of light across space.]

Photoelectricity.—HERTZ and others in their experiments on the discharge of electricity through gases noticed that the discharge between two terminals took place more easily when the spark gap was illuminated. In this respect ultra-violet light was more effective than rays from the visible or infra-red regions of the spectrum. HALLWACHS soon afterwards showed that this phenomenon depended on illumination of the cathode. Further investigations by ELSTER and GEITEL, and by LENARD, proved that when metals are illuminated they emit electrons—the so-called ***photo-electrons.***

Experiment.—Connect a zinc plate to an electroscope and charge it negatively. Then allow light from an arc lamp (rich in ultra-violet rays) to fall on the plate. Its potential diminishes rapidly, due to the escape of electrons from its surface.

Repeat the above with the plate charged positively. Its potential does not change since the positive charge on the plate prevents the electrons from escaping.

A more exact study of the photoelectric effect may be made with the apparatus shown in Fig. 51·13. The plates A and B are

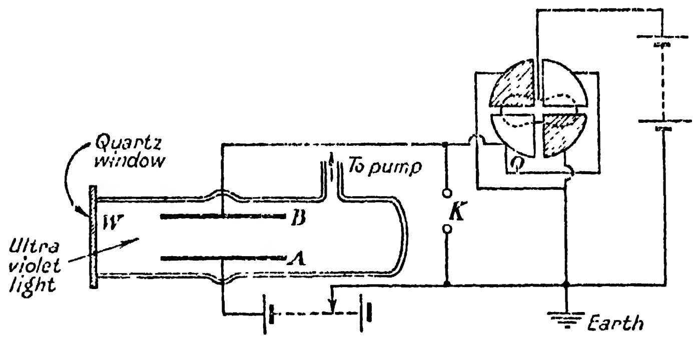


FIG. 51·13.—Photo-electric Currents.

enclosed in an exhausted glass tube fitted with a quartz window W and connected to the positive pole of a high-tension battery and to the insulated quadrants of an electrometer, Q, respectively. The negative pole of the battery is earthed. The key, K, enables the electrometer to be discharged before commencing the experiment. When this is removed and ultra-violet light allowed to fall on B the electrometer needle begins to move, showing that electrons are emitted from B.

It is now known that no photoelectrons are emitted unless the frequency of the incident radiation is greater than a certain value characteristic for each metal. Thus for sodium there is no photoelectric effect unless the incident light has a frequency greater than

about 5×10^{14} (green light). Thus blue light (shorter wave-length and therefore greater frequency than green light) falling on sodium causes the emission of many photoelectrons. If red light, however, is used no such emission occurs even if the light is incident for years—theory suggests it to be impossible.

Three important generalizations with reference to the emission of photoelectrons are as follows :—

(*a*) The number of electrons emitted per second is directly proportional to the intensity of the incident radiation.

(*b*) The kinetic energy of the electrons is independent of the intensity of the light.

(*c*) The kinetic energy of the electrons increases with the frequency of the incident light according to the following law due to EINSTEIN :—If m is the mass and v the velocity of an electron, ν the frequency of the incident light, and ν_0 the characteristic frequency for the particular metal under investigation, then

$$\tfrac{1}{2}mv^2 = h(\nu - \nu_0),$$

where h is a universal constant, termed Planck's constant. It is equal to $6{\cdot}55 \times 10^{-27}$ erg. sec.

EXAMPLES LI

1. Give a sketch of two types of X-ray bulb and explain how the rays are produced. What are the chief properties of X-rays ?

2. Describe how cathode rays and Röntgen rays may be produced. What are the essential differences between these two types of rays ?

3. Give a short account of the more important emanations from a radio-active substance.

4.—Electrons enter a uniform magnetic field of intensity 20π oersted in a direction at right angles to the lines of force. In what time do the electrons describe a complete circle ?

5.—A beam of electrons is emitted from a hot filament in an easterly direction. The horizontal component of the earth's magnetic field (0·2 oersted) deflects the beam into a circular arc of radius 2 metres. Calculate a value for the velocity of the electrons. Through what p.d. must the electrons fall in order to acquire this velocity, assuming that the velocity of escape is zero ?

6.—A drop of oil, density 2·0 gm. cm.$^{-3}$ and radius 0·0001 cm., carries a charge of four electrons. What p.d. must be applied between the plates of the condenser in Millikan's experiment in order that the drop may float if the plates are 5 mm. apart ? Also calculate the maximum rate of fall of the drop when the electric field is removed, if the viscosity of air at 15° C., the temperature at which the experiment is carried out, is $1{\cdot}80 \times 10^{-4}$ gm. cm.$^{-1}$ sec.$^{-1}$

ANSWERS TO THE EXAMPLES

I. (1) 0·82, 0·73, 0·68, 1·07, 57° 18′. (4) 3·14 ft. (5) 24·9 cm.

II. (1) 41·7 ft. (2) 4° 47′. (3) 0·92 ft. sec.$^{-2}$, 6·0 sec. (4) 8 cm. sec.$^{-2}$ (5) 4·79 sec., 367 ft. (6) 2·7 sec. (7) 48 ml. hr.$^{-1}$ (8) 67·9 ft. sec.$^{-1}$, 102 ft. (9) 20 cm., 40 cm. (10) 110 ft. lb.-wt. (11) 0·447, 1. (12) $20\pi/3$, $40\pi^2/9$. (13) 1·01 ton.-wt. (14) $2\pi/7$ sec. (15) $4{\cdot}3 \times 10^6$ erg. (16) $10\,m\sqrt{lg(2-\sqrt{3})}$ gm. cm. sec.$^{-1}$, $50\,mlg(2-\sqrt{3})$ erg.

III. (1) 139 ft., 324 lb. ft. sec.$^{-1}$ (2) 11·7 lb.-wt., 3·34 ft. sec.$^{-2}$ (4) 0·29 cm. from the centre. (5) 42·2 lb.-wt. (6) 28 lb.-wt., 1 : 3. (8) 20·67 gm. (9) 1·55 ft. from the fulcrum. (10) 78·7 gm. (11) 62° 55′. (13) Any point in a vertical line 0·5 in. from the median through *C*, and on the side nearer to *B*. (17) 86 ft. (18) 16·8 ft. (19) 26′.

IV. (1) 20·4 gm. (2) 40·6 cm. (3) 160·3 atmos., 152·1 ton.-wt. ft.$^{-2}$ (4) 211 gm.-wt. cm.$^{-2}$ (5) 0·73 ft. (6) 6·88 in. (10) 4·95 cm.^{3} (11) 29·2 in. of mercury. (12) 261·1 gm. (13) 0·0779 cm., 1·297 gm. cm.$^{-3}$ (14) 13·7 cm. of water. (15) 12·5 cm.^{3} (16) 3,809 lb. yd.$^{-3}$ (18) 0·604 : 1 by volume. (19) 0·0169 gm. too heavy. (20) 7·57 cm. (21) 1·014. (22) 75 cm. (23) 57 cm. (25) 29·54 in. (26) 29·5 in.

V. (3) 30·5 dyne. cm.$^{-1}$ (4) 2·7 gm. cm.$^{-1}$ sec.$^{-1}$ (8) 30·7 dyne. cm.$^{-1}$ (13) $1{\cdot}25 \times 10^5$ dyne. cm.$^{-2}$ (15) 2·50 cm. (16) 1·3 cm. (17) 15·1 cm.

VI. (1) 2. (2) 311 lb.-wt. in.$^{-2}$ (3) 8 kgm. (4) 0·085 cm.^{3} (5) $6{\cdot}68 \times 10^4$ erg. (6) $2{\cdot}03 \times 10^{12}$ dyne. cm.$^{-2}$ (7) 32 kgm. (9) $4{\cdot}05 \times 10^{12}$ dyne. cm.$^{-2}$ (10) 1·0007 gm. cm.$^{-3}$ (11) $1{\cdot}29 \times 10^4$ lb.-wt. in.$^{-2}$ (12) 1·76 sec. (14) 179·5 cm.

VII. (1) 99·59° C., + 0·39° C.

VIII. (1) 0·0000075 deg.$^{-1}$ C. (3) 0·9998 sec. (4) $3{\cdot}82 \times 10^{-4}$ deg.$^{-1}$ C. $3{\cdot}77 \times 10^{-4}$ deg.$^{-1}$ C.

IX. (1) 13·35 gm. cm.$^{-3}$ (2) 0·00014 deg.$^{-1}$ C. (3) 0·0664 cm.^{3} at 0° C. (4) 241·5 cm.^{3}, 420 cm.^{3} (5) 83° C. (6) 80·7 cm. of mercury. (8) $1{\cdot}064 \times 10^{-4}$ deg.$^{-1}$ C. (9) 0·00018 deg.$^{-1}$ C. (10) 0·340 gm. (13) 0·00215 deg.$^{-1}$ F., or 0·00387 deg.$^{-1}$ C., i.e. 1·06 times too large. (14) 0·128 of the internal volume. (16) 0·27 cm. (17) 321·5° C. (19) 0·00072. (21) 751·7 mm. (22) 0·00031 deg.$^{-1}$ C. (24) 225° C. (25) $2{\cdot}29 \times 10^{-4}$ deg.$^{-1}$ C.

X. (1) $6{\cdot}74 \times 10^3$ cal. (2) 13·4 cal. deg.$^{-1}$ C. (3) 0·092 cal. gm.$^{-1}$ deg.$^{-1}$ C. (4) 11·2 gm. (5) 0·141 cal. gm.$^{-1}$ deg.$^{-1}$ C. (6) 0·077 cal. gm.$^{-1}$ deg.$^{-1}$ C. (7) 0·045 cal. gm.$^{-1}$ deg.$^{-1}$ C. (14) 0·072 cal. gm.$^{-1}$ deg.$^{-1}$ C., 0·63 cal. deg.$^{-1}$ C. (15) 138·9 cal. deg.$^{-1}$ C. (16) 62 lb.

XII. (1) 7·6 cm. of mercury. (2) 0·42 atmos. (3) 0·8 and 1·2 cm. of mercury respectively. (4) 23·7 cm.

XIV. (1) 71 per cent. (2) 349 metre. sec.$^{-1}$ (3) $2{\cdot}11 \times 10^3$ cm. sec.$^{-1}$ (5) 0·0088° C.

XV. (1) 0·118 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C. (3) 2×10^6 cal. (4) $1{\cdot}67 \times 10^4$ cal. sec.$^{-1}$ (8) $1{\cdot}5 \times 10^{-4}$ cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C. (9) 0·93 cal. cm.$^{-1}$ sec.$^{-1}$ deg.$^{-1}$ C. (10) 464 watt.

XVI. (3) $3{\cdot}83 \times 10^{-5}$ cal. cm.$^{-2}$ sec.$^{-1}$ deg.$^{-1}$ C. (4) 0·17 cal. sec.$^{-1}$ (5) 73·1 cal. sec.$^{-1}$ (6) 8 : 1, equal. (8) 3·71 deg. C. min.$^{-1}$ (9) Temp. rises at a constant rate 0·05 deg. C. sec.$^{-1}$ (11) $5{\cdot}9 \times 10^{-3}$ deg. C. sec.$^{-1}$ (12) $\dfrac{4\sigma A T_0^3}{Jms}\left[\phi + \tfrac{3}{2}\dfrac{\phi^2}{T_0} + \dfrac{\phi^3}{T_0^2} + \tfrac{1}{4}\dfrac{\phi^4}{T_0^3}\right]$. (13) 20° C., 44° C. (14) 3·11 min.

XVII. (1) 1 : 1·82. (2) 58·1 cm. (3) 1·04 (or 0·96). (4) 14·3 cm. (5) 1·06 or 18·94 metres from the brighter source. (6) 41·9 cm. from spot.

XVIII. (1) 62°.

XIX. (1) 12 cm. (2) — 11·3 cm. (3) 2·7 ft. in front of mirror. (4) 3·26 in. behind convex mirror. (5) —2·2 in. (6) —5·2 cm. (7) 52·5 cm. (8) 2·67 ft.

XX. (2) 28° 15′. (3) 1·61. (4) 48° 33′, 62° 44′. (5) 1·605, 38° 33′. (6) 1·529, 39° 44′. (8) 0·38 in. (9) 33·2 cm. behind the front surface. (10) 34° 51′. (11) 4·81 cm. (12) 68° 38′. (13) 69° 42′.

XXI. (1) — 22·3 cm. (2) + 8·4 cm., 0·41. (3) 41·4 cm. (4) 20 cm. behind second lens ; unity. (5) — 17·7 cm. (6) 41·0 cm. (7) 3·2 cm. from

mirror; 0·92 cm. (10) 20·6 cm. (11) 1·53. (12) 2·5 cm. from the surface. (13) 35 cm., 25 cm., 10·4 cm., 1·51.

XXIII. (6) + 10 cm., − 6·7 cm. (7) $f_A = -37·5$ cm., $f_B = +60$ cm. (8) 19·8 cm., 19·8 cm., − 408 cm. (9) 57·5 cm., 42·6 cm. *or* 24·4 cm., 57·5 cm.

XXIV. (1) − 5·4 in. (2) 5 ft. (3) 24 in.; 1·64. (4) − 39·4 in. (5) 4·2 cm., 6.

XXV. (3) 10·5 in., 4·9 ft. square. (7) 9·7. (10) 3·69 cm., 14·4.

XXIX. (1) $6·87 \times 10^{-5}$ cm. (2) $2·2 \times 10^{-5}$ radian. (7) 341 cm. (8) $5·73 \times 10^{-5}$ cm. (9) $\theta_0 = -45°$, $\theta_1 = -20°\ 19'$, $\theta_2 = 0°\ 44'$, $\theta_3 = 21°\ 54'$, $\theta_4 = 47°\ 8'$ and all negative orders are absent.
(10) (*a*) $1·09 \times 10^{-2}$ cm., (*b*) $0·55 \times 10^{-2}$ cm. (11) 0·367 cm.; ring system has a bright centre. (12) 0·020 cm. (13) 1·46 cm. (14) 1′, 3′. (15) 1·72, 1·60 cm. on one side of the green line, 5·71 cm. and 5·80 cm. on the other side.

XXX. (1) 1·624.

XXXII. (1) 494 cycle. sec.$^{-1}$ (2) $\frac{1}{4000}$ c, where c = velocity of light. (3) 559 cycle. sec.$^{-1}$, 469 cycle. sec.$^{-1}$

XXXIII. (1) (*a*) 16·7, 50·0, 83·3, 116·7, 150 cm. (*b*) 17·2, 51·5, 85·9, 120·3 cm. (3) 8·3 gm. cm.$^{-3}$ (4) Increase tension 2·78 times; reduce length 0·60 times. (5) 6·4. (10) $3·46 \times 10^4$ cm. sec.$^{-1}$ (11) 469·3 cycle. sec.$^{-1}$ (12) 10 kgm.-wt. (13) $f/f_T = 10\sqrt{10}$. (14) ($256 \pm \frac{1}{12}$) cycle. sec.$^{-1}$ (15) 150·2 cycle. sec.$^{-1}$

XXXV. (2) 28·6 e.s.u. (3) 0·30 dyne. (4) 0·92 dyne.

XXXVI. (1) 3·37 cm., 5·91 e.s.u. (2) 7·5 e.s.u., 2250 volt. (3) 28 cm. (4) 0·79 e.s.u. less. (7) 0·040 e.s.u., 0·10 erg. (9) 31·4 erg. (16) $1·25 \times 10^5$ erg., 40 volt., $1·00 \times 10^5$ erg. (18) 28·3, 20, 20; 8·33, 0, 0 e.s.u.

XXXVII. (2) 1·77 dyne., 6·19 dyne. (3) − 2·65 e.s.u. of charge, $4·42 \times 10^{-7}$ dyne. cm.$^{-2}$ (4) $\kappa \div 2\pi t(\kappa + 1)$. (5) 6π dyne. cm.$^{-2}$ (6) $8·3 \times 10^{-6}$ cm.

XXXVIII. (3) $1·62 \times 10^{-10}$ amp. (4) $3·1 \times 10^2$ dyne.

XXXIX. (1) − 4·7 dyne. (2) 0·71 dyne. (3) 2·99 cm. (4) 63·5 or 31·7 erg. oersted.$^{-1}$ (8) 90 erg. oersted.$^{-1}$

XL. (1) 0·0225 oersted. (2) 0·14 oersted. (3) 37·2 unit-poles. 744 unit-pole. cm. (4) 9·7 dyne. cm. (5) 7·4 sec. (6) 73·9 unit-pole. cm. (7) 125·3 dyne. cm., 65·2 erg. (15) 8·5. (16) 158 erg. oersted.$^{-1}$ (17) 22·7 unit-pole.

XLI. (2) 229 erg. oersted.$^{-1}$, 0·17 oersted.

XLIV. (4) 0·159 amp. (5) $19·98\pi$ oersted. (6) $0·4r^{-1}$ oersted. (7) 14·1 cycle. min.$^{-1}$

XLV. (1) 110·2 volt. (2) 6·76 volt. (3) 0·097 amp. (4) 1·76 volt., 1·62 volt. (5) 3·6 ohm. (6) 9 ohm., 0·92 ohm., 0·23 amp., 2·25 amp. (7) 10 ohm. (8) 1101 ohm. (9) 118·2 ohm. (10) 1·43 ohm., 2 ohm. (11) 0·16 oersted. (12) 2·18 volt. (13) 1 : 1·83. (14) 6 ohm., $18·8 \times 10^{-6}$ ohm. cm. (16) 0·24 amp. (17) 4·55 volt.; 10 volt. (18) 9,000 ohm. series resistance. (19) 9,900 ohm. series resistance. (20) 6·90 ohm. (22) 19·1 ohm.; 1·01. (23) Shunt with 10·1 ohm.

XLVI. (1) 1·984. (2) 1·51. (3) 1·015 amp. (6) $1·77 \times 10^{-6}$ ohm. cm. (10) 21·3 erg. (11) 6 ohm., 4·55 cm. (12) 50 divisions, 2×10^4 ohm.

XLVII. (1) 0·28 gm. (2) 8 turns. (3) 2·15 hour. (6) 13 ohm. (7) 1·25 amp. (9) 1·15 amp. 0·068 gm. (16) $2·69 \times 10^{19}$. (17) $1·60 \times 10^{-19}$ coulomb.

XLVIII. (1) 4·08. (2) 0·98. (3) 4·7 ohm. (6) 100 : 10 : 18·2 : 1·82 or 1 : 10 : 5·5 : 55. (7) Maximum output when R = 2 ohm. = internal resistance of battery. (9) 7·23 ohm.

XLIX. (5) $7·1 \times 10^{-5}$ volt. (6) 0·050 amp. (7) 4×10^{-4} volt. (8) $1·3 \times 10^{-3}$ volt. (9) $1·3 \times 10^{-4}$ volt. (13) 2×10^3 oersted. (14) $4\pi\mu\lambda^{-1}n^2\alpha$. (15) 2×10^{-4} henry. (16) $5\pi \times 10^{-4}$ volt. (17) $9·5 \times 10^4$ linkages.

L. (2) 60, 4·7. (3) 26·3 unit-pole. (4) 0·195 oersted. (5) $2·37 \times 10^3$ gauss. (6) 180 gauss. (7) 18·4 unit-pole., 368 gauss, $4·62 \times 10^3$ gauss, $0·82 \times 10^3$.

LI. (4) $0·56 \times 10^{-8}$ sec. (5) $7·08 \times 10^6$ cm. sec.$^{-1}$, 139 volt. (6) $6·46 \times 10^2$ volt.; $2·4 \times 10^{-2}$ cm. sec.$^{-1}$

TRIGONOMETRICAL RATIOS

Angle.									
Degrees.	Radians.	Chord.	Sine.	Tangent.	Co-tangent.	Cosine.			
0°	0	0	0	0	∞	1	1·414	1·5708	90°
1	·0175	·017	·0175	·0175	57·2900	·9998	1·402	1·5533	89
2	·0349	·035	·0349	·0349	28·6363	·9994	1·389	1·5359	88
3	·0524	·052	·0523	·0524	19·0811	·9986	1·377	1·5184	87
4	·0698	·070	·0698	·0699	14·3007	·9976	1·364	1·5010	86
5	·0873	·087	·0872	·0875	11·4301	·9962	1·351	1·4835	85
6	·1047	·105	·1045	·1051	9·5144	·9945	1·338	1·4661	84
7	·1222	·122	·1219	·1228	8·1443	·9925	1·325	1·4486	83
8	·1396	·140	·1392	·1405	7·1154	·9903	1·312	1·4312	82
9	·1571	·157	·1564	·1584	6·3138	·9877	1·299	1·4137	81
10	·1745	·174	·1736	·1763	5·6713	·9848	1·286	1·3963	80
11	·1920	·192	·1908	·1944	5·1446	·9816	1·272	1·3788	79
12	·2094	·209	·2079	·2126	4·7046	·9781	1·259	1·3614	78
13	·2269	·226	·2250	·2309	4·3315	·9744	1·245	1·3439	77
14	·2443	·244	·2419	·2493	4·0108	·9703	1·231	1·3265	76
15	·2618	·261	·2588	·2679	3·7321	·9659	1·218	1·3090	75
16	·2793	·278	·2756	·2867	3·4874	·9613	1·204	1·2915	74
17	·2967	·296	·2924	·3057	3·2709	·9563	1·190	1·2741	73
18	·3142	·313	·3090	·3249	3·0777	·9511	1·176	1·2566	72
19	·3316	·330	·3256	·3443	2·9042	·9455	1·161	1·2392	71
20	·3491	·347	·3420	·3640	2·7475	·9397	1·147	1·2217	70
21	·3665	·364	·3584	·3839	2·6051	·9336	1·133	1·2043	69
22	·3840	·382	·3746	·4040	2·4751	·9272	1·118	1·1868	68
23	·4014	·399	·3907	·4245	2·3559	·9205	1·104	1·1694	67
24	·4189	·416	·4067	·4452	2·2460	·9135	1·089	1·1519	66
25	·4363	·433	·4226	·4663	2·1445	·9063	1·075	1·1345	65
26	·4538	·450	·4384	·4877	2·0503	·8988	1·060	1·1170	64
27	·4712	·467	·4540	·5095	1·9626	·8910	1·045	1·0996	63
28	·4887	·484	·4695	·5317	1·8807	·8829	1·030	1·0821	62
29	·5061	·501	·4848	·5543	1·8040	·8746	1·015	1·0647	61
30	·5236	·518	·5000	·5774	1·7321	·8660	1·000	1·0472	60
31	·5411	·534	·5150	·6009	1·6643	·8572	·985	1·0297	59
32	·5585	·551	·5299	·6249	1·6003	·8480	·970	1·0123	58
33	·5760	·568	·5446	·6494	1·5399	·8387	·954	·9948	57
34	·5934	·585	·5592	·6745	1·4826	·8290	·939	·9774	56
35	·6109	·601	·5736	·7002	1·4281	·8192	·923	·9599	55
36	·6283	·618	·5878	·7265	1·3764	·8090	·908	·9425	54
37	·6458	·635	·6018	·7536	1·3270	·7986	·892	·9250	53
38	·6632	·651	·6157	·7813	1·2799	·7880	·877	·9076	52
39	·6807	·668	·6293	·8098	1·2349	·7771	·861	·8901	51
40	·6981	·684	·6428	·8391	1·1918	·7660	·845	·8727	50
41	·7156	·700	·6561	·8693	1·1504	·7547	·829	·8552	49
42	·7330	·717	·6691	·9004	1·1106	7431	·813	·8378	48
43	·7505	·733	·6820	·9325	1·0724	·7314	·797	·8203	47
44	·7679	·749	·6947	·9657	1·0355	·7193	·781	·8029	46
45°	·7854	·765	·7071	1·0000	1·0000	·7071	·765	·7854	45°
			Cosine.	Co-tangent.	Tangent.	Sine.	Chord.	Radians.	Degrees.
								Angle.	

	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170						4	9	13	17	21	26	30	34	38
						0212	0253	0294	0334	0374	4	8	12	16	20	24	28	32	37
11	0414	0453	0492	0531	0569						4	8	12	15	19	23	27	31	35
						0607	0645	0682	0719	0755	4	7	11	15	19	22	26	30	33
12	0792	0828	0864	0899	0934	0969					3	7	11	14	18	21	25	28	32
							1004	1038	1072	1106	3	7	10	14	17	20	24	27	31
13	1139	1173	1206	1239	1271						3	7	10	13	16	20	23	26	30
						1303	1335	1367	1399	1430	3	7	10	12	16	19	22	25	29
14	1461	1492	1523	1553							3	6	9	12	15	18	21	24	28
					1584	1614	1644	1673	1703	1732	3	6	9	12	15	17	20	23	26
15	1761	1790	1818	1847	1875	1903					3	6	9	11	14	17	20	23	26
							1931	1959	1987	2014	3	5	8	11	14	16	19	22	25
16	2041	2068	2095	2122	2148						3	5	8	11	14	16	19	22	24
						2175	2201	2227	2253	2279	3	5	8	10	13	15	18	21	23
17	2304	2330	2355	2380	2405	2430					3	5	8	10	13	15	18	20	23
							2455	2480	2504	2529	2	5	7	10	12	15	17	19	22
18	2553	2577	2601	2625	2648						2	5	7	9	12	14	16	19	21
						2672	2695	2718	2742	2765	2	5	7	9	11	14	16	18	21
19	2788	2810	2833	2856	2878						2	4	7	9	11	13	16	18	20
						2900	2923	2945	2967	2989	2	4	6	8	11	13	15	17	19
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201	2	4	6	8	11	13	15	17	19
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2	4	6	8	10	12	14	16	18
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2	4	6	8	10	12	14	15	17
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2	4	6	7	9	11	13	15	17
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2	4	5	7	9	11	12	14	16
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2	3	5	7	9	10	12	14	15
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2	3	5	7	8	10	11	13	15
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2	3	5	6	8	9	11	13	14
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2	3	5	6	8	9	11	12	14
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1	3	4	6	7	9	10	12	13
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1	3	4	6	7	9	10	11	13
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1	3	4	6	7	8	10	11	12
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1	3	4	5	7	8	9	11	12
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1	3	4	5	6	8	9	10	12
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1	3	4	5	6	8	9	10	11
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1	2	4	5	6	7	9	10	11
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1	2	4	5	6	7	8	10	11
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1	2	3	5	6	7	8	9	10
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1	2	3	5	6	7	8	9	10
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1	2	3	4	5	7	8	9	10
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1	2	3	4	5	6	8	9	10
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1	2	3	4	5	6	7	8	9
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1	2	3	4	5	6	7	8	9
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1	2	3	4	5	6	7	8	9
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1	2	3	4	5	6	7	8	9
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1	2	3	4	5	6	7	8	9
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1	2	3	4	5	6	7	7	8
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1	2	3	4	5	5	6	7	8
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1	2	3	4	4	5	6	7	8
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1	2	3	4	4	5	6	7	8
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1	2	3	3	4	5	6	7	8

	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1	2	3	3	4	5	6	7	8
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1	2	2	3	4	5	6	7	7
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1	2	2	3	4	5	6	6	7
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1	2	2	3	4	5	6	6	7
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1	2	2	3	4	5	5	6	7
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1	2	2	3	4	5	5	6	7
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1	2	2	3	4	5	5	6	7
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1	1	2	3	4	4	5	6	7
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1	1	2	3	4	4	5	6	7
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1	1	2	3	4	4	5	6	6
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1	1	2	3	4	4	5	6	6
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1	1	2	3	3	4	5	6	6
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1	1	2	3	3	4	5	5	6
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1	1	2	3	3	4	5	5	6
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1	1	2	3	3	4	5	5	6
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1	1	2	3	3	4	5	5	6
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1	1	2	3	3	4	5	5	6
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1	1	2	3	3	4	4	5	6
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1	1	2	2	3	4	4	5	6
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1	1	2	2	3	4	4	5	6
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1	1	2	2	3	4	4	5	5
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1	1	2	2	3	4	4	5	5
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1	1	2	2	3	4	4	5	5
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1	1	2	2	3	4	4	5	5
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1	1	2	2	3	3	4	5	5
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859	1	1	2	2	3	3	4	5	5
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1	1	2	2	3	3	4	4	5
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1	1	2	2	3	3	4	4	5
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1	1	2	2	3	3	4	4	5
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1	1	2	2	3	3	4	4	5
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1	1	2	2	3	3	4	4	5
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1	1	2	2	3	3	4	4	5
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1	1	2	2	3	3	4	4	5
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1	1	2	2	3	3	4	4	5
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1	1	2	2	3	3	4	4	5
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1	1	2	2	3	3	4	4	5
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	0	1	1	2	2	3	3	4	4
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	0	1	1	2	2	3	3	4	4
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	0	1	1	2	2	3	3	4	4
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0	1	1	2	2	3	3	4	4
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0	1	1	2	2	3	3	4	4
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0	1	1	2	2	3	3	4	4
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0	1	1	2	2	3	3	4	4
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0	1	1	2	2	3	3	4	4
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0	1	1	2	2	3	3	4	4
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0	1	1	2	2	3	3	4	4
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0	1	1	2	2	3	3	4	4
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0	1	1	2	2	3	3	4	4
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0	1	1	2	2	3	3	3	4

INDEX